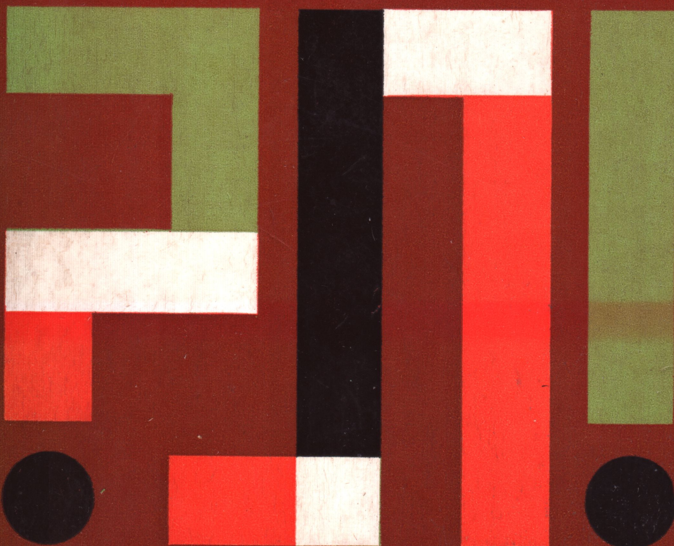
THEIL/BOOT/KLOEK



VOORSPELLEN EN BESLISSEN

INLEIDING TOT DE KWANTITATIEVE ECONOMIE EN OPERATIONELE RESEARCH



CVD

De toekomst is onzeker. Wanneer zal de machine defect raken? Hoeveel zal ik de volgende maand verkopen? Hoe beinvloeden de belastingtarieven het investeringsniveau? Temidden van deze onzekerheid moeten beslissingen genomen worden, die goed of slecht blijken te zijn al naar gelang de toekomst zus of zo uitvalt. Men kan het probleem niet uit de weg gaan. Géén beslissing is óók een beslissing!
Het tweede deel van dit boek (vanaf hoofdstuk 6) handelt over de vraag hoe men 'juiste' beslissingen moet nemen in zulke onzekere situaties. Hoewel een *werkelijk* juiste beslissing (dat is een beslissing die u, ook achteraf beschouwd, precies zo zou hebben genomen) geen haalbare kaart is, kan een nauwgezette bestudering van het probleem in belangrijke mate bijdragen tot een rationele besluitvorming. Die nauwgezette bestudering vereist een inzicht in de samenhang van het economisch gebeuren — daarover handelen verscheidene hoofdstukken van het eerste deel.
In de meeste gevallen komt de doelstelling neer op het minimaliseren van kosten of tijdverlies of afstand of risico; of op het zo groot mogelijk maken van winst of het nationale inkomen enz. Dat minimaliseren of maximaliseren is een wiskundige opgaaf. Ook het weergeven van samenhangen is in vele gevallen een probleem met mathematische aspecten. Een woord van geruststelling: de gebruikte wiskunde kan redelijkerwijs voor niemand met een middelbare opleiding een struikelblok vormen dit boek te lezen. De toon van het boek, dat niet pretendeert als substituut voor een zesjarige universitaire opleiding te dienen, is bepaald luchthartig. Desondanks hopen de auteurs erin geslaagd te zijn u een inzicht te geven in de gedachtengang en in het belang van de moderne theorie van de rationele besluitvorming.

MARKA-BOEKEN

Prof. dr. H. Theil / Drs. J. C. G. Boot / Drs. T. Kloek

VOORSPELLEN EN BESLISSEN



HET POCKETBOEK

VOOR ORGANISATIE EN BEDRIJF

ONDER HOOFDREDACTIE VAN

DR. IR. H. DE BOER

UTRECHT

MARKA-BOEKEN

ANTWERPEN

PROF. DR. H. THEIL / DRS. J. C. G. BOOT
DRS. T. KLOEK

VOORSPELLEN EN BESLISSEN

INLEIDING TOT DE KWANTITATIEVE ECONOMIE EN OPERATIONELE RESEARCH

UTRECHT
MARKA-BOEKEN
ANTWERPEN

Prisma-boeken, Prisma-compendia, Prisma-juniores, Prisma-kinderpockets, Aula-, Marka- en Pictura-boeken worden uitgegeven door Uitgeverij Het Spectrum N.V., Utrecht/Antwerpen.

INHOUD

PROLOOG 9

1. LINEAIR PROGRAMMEREN
   1. De structuur van een lineair programmeringsprobleem 14
   2. Productie van radio's 18
   3. Een grafische afbeelding 19
   4. De oplossing bij verschillende winstmarges 25
   5. Spelingsvariabelen 27
   6. De oplossingsfilosofie van de Simplex-techniek 30
   7. Wat is de prijs van een nevenvoorwaarde? 33
   8. Wedden op Duindigt 36
   9. Een fabrikant 39
   Literatuur 47

2. HET OPTIMALE EN HET KRITIEKE PAD
   1. Brand 49
   2. Van noord-west naar zuid-oost 53
   3. Wanneer is een transportplan optimaal? 57
   4. Hoe maken we een niet-optimaal transportplan optimaal? 62
   5. Een quiz 67
   6. De gedachtengang van het kritieke pad 69
   7. Het kritieke pad van dit boek 72
   Literatuur 75

3. AAN- EN AFVOERANALYSE
   1. Samenhang in de economie 77
   2. Aan- en afvoertabellen in de practijk 81
   3. Doel en veronderstellingen van de aan- en afvoeranalyse 85
   4. Hoe werkt de aan- en afvoeranalyse? 88
   5. Nog enkele details 93
   6. Knelpunten 98
   Literatuur 100

4. ECONOMETRISCHE MACRO-MODELLEN
   1. Nogmaals: samenhang in de economie 101

2. Een heel eenvoudig model 105
3. Een klein model van de Verenigde Staten 108
4. Nogmaals: het kleine model van de Verenigde Staten 117
5. Over gevolgen op langere termijn 118
6. Grote modellen 120
Literatuur 124

5. OVER ECONOMISCHE VOORSPELLINGEN
1. Voorspellen: kunst of wetenschap? 126
2. De investeringsenquête van het C.B.S. 131
3. Econometrische modelvoorspellingen 139
4. De voorspellingskwaliteit van de Centrale Economische Plannen 142
5. Aan- en afvoervoorspellingen 147
Literatuur 154

6. OVER ONZEKERHEID EN WAARSCHIJNLIJKHEID
1. Onzekerheid: een bijna universeel verschijnsel 157
2. De gedachtengang van de waarschijnlijkheidsrekening 158
3. Discrete verdelingen 164
4. Verwachting en variantie 166
5. Het rekenen met verwachtingen 169
6. Over continue verdelingen 171
Literatuur 180
Aanhangsel 180

7. HET STRATEGIEBEGRIP
1. Strategie en tactiek 185
2. Voordelige geldbelegging 187
3. De waardering van de resultaten 190
4. Vier paren van opeenvolgende beslissingen 194
5. Vier strategieën 198
Literatuur 203

8. SPELTHEORIE
1. Twee luchtvaartmaatschappijen 204
2. Spel, zet en strategie 206
3. Minimax en zadelpunt 210
4. Gemengde strategieën 214
5. Een grafische oplossing 217

6. De minimax-stelling 219
7. Over meer spelers en niet-nulspelen 224
Literatuur 226

9. WACHTEN
1. Het probleem 227
2. Aankomst 229
3. Bediening 233
4. De gemiddelde rijlengte 235
5. Voorrang 238
6. Een machinehal 242
Literatuur 245

10. SIMULATIE EN BELEIDSSPELEN
1. De voetbalpool 246
2. Een margarinefabriek 248
3. Verfijningen 255
4. Schaak en monopoly 257
5. Bromfietsen 259
6. Problemen bij beleidsspelen 262
7. In de directiekamer 265
Literatuur 267

11. OVER PRODUCTIE- EN VOORRAADBESLISSINGEN
1. De optimale seriegrootte 268
2. Gevoeligheidsanalyse 274
3. Een uitvoeriger kostenfunctie 277
4. Het probleem van de kostenminimering 285
5. Lineaire beslissingsregels 290
6. Concrete gevallen 295
7. Voorspellingen van toekomstige beslissingen 301
Literatuur 305

12. DE STATISTISCHE SPECIFICATIE VAN ECONOMISCHE RELATIES
1. Het probleem; de twee bronnen van informatie 307
2. De puntenwolk 310
3. Aanpassing volgens kleinste kwadraten 313
4. De gedachtengang van het statistisch schatten 319
5. Zuivere schattingen 324
6. Schatting in de regressieanalyse 326

7. Standaardfouten 334
Literatuur 340
Aanhangsel 340

13. OM DE GULDEN VAN DE CONSUMENT
1. Waarom is het consumentengedrag interessant? 344
2. Reëel inkomen en relatieve prijzen 345
3. Hoeveel wordt er gekocht? 348
4. Inkomenselasticiteiten 351
5. Prijselasticiteiten 353
6. Huishoudrekeningen 355
7. Kwaliteit en kwantiteit 359
8. Andere factoren waardoor de consument zich laat leiden 361
Literatuur 363

EPILOOG 365

REGISTER 371

PROLOOG

De voortdurend belangrijker wordende rol van de wiskunde is een interessant aspect van de hedendaagse geschiedenis van de wetenschap. Met niet al te veel overdrijving kan men stellen, dat de beoefening van de wiskunde vroeger een in zichzelf gesloten systeem was. Het vak was er om op de middelbare school bestudeerd te worden; en als men er dieper in wilde doordringen betekende dit maatschappelijk bezien in de meeste gevallen een voorbereiding op het leraarsambt, dus op het doceren van dezelfde stof in de middelbare school opdat de volgende generatie met het vak kon kennismaken. Maar tegenwoordig worden wij geconfronteerd met kosmonauten en astronauten die door het luchtruim snellen op koersen die tot op seconden nauwkeurig berekend dienen te worden. Dat gaat niet zonder wiskunde. In het bijzonder, dat gaat niet zonder electronische rekenmachines – één van de spectaculaire vindingen van de laatste tijd. We zijn aan het bestaan van dergelijke machines al zozeer gewend, dat we er in feite maar nauwelijks bij stil staan dat de eerste commerciële exemplaren pas in het midden van de vijftiger jaren in ons land zijn ingevoerd. Eigenlijk zitten we dus nog maar in het begin van de ontwikkeling op dit terrein. Dit wordt op treffende wijze geïllustreerd door een recente studie van de Stichting Studiecentrum voor Administratieve Automatisering, waaruit blijkt dat Nederland en de omringende landen rond vijf jaren bij de Verenigde Staten ten achter liggen, dat binnen een beperkt aantal jaren hier een vertienvoudiging van het aantal machines valt te verwachten, en dat desondanks de achterstand t.o.v. de Verenigde Staten dan niet zal zijn ingelopen. Integendeel, die achterstand wordt dan verwacht tot ongeveer acht jaren toe te nemen, omdat de ontwikkeling daar nog weer sneller is dan hier.

De rekenapparatuur is belangrijk voor de toepassing van mathematische technieken, maar minstens even belangrijk is de ontwikkeling van die technieken zelf. Vroeger beperkte de toepassing van de wiskunde zich in hoofdzaak tot natuur- en technische wetenschappen; de economie is daar sinds enkele decennia bijgekomen. Voorlopers kan men al in veel vroegere perioden aanwijzen, tot in het midden van de negentiende eeuw toe; maar de ontwikkeling kwam pas goed op gang toen in het

begin van de dertiger jaren de Econometric Society werd opgericht. Na de Tweede Wereldoorlog is als jonger zusje van de econometrie de 'operations research' tot ontplooiing gekomen; in het Nederlands wordt deze term meestal als operationele research vertaald, maar sommigen vinden dat te letterlijk en prefereren besliskunde of bedrijfseconometrie. Dit vak houdt zich i.h.b. bezig met de mathematische zijde van bedrijfseconomische en bedrijfsorganisatorische vraagstukken. De belangrijkste stimulans van dit vak was de oorlog zelf. Toen nl. Engeland na Duinkerken de strijd alleen had te voeren, viel het besluit de wetenschap te mobiliseren om de problemen van de oorlog op de meest doelmatige manier aan te pakken: scheepstransporten, duikbootafweer, bombardementen, enz. Men bracht daartoe een groep van wetenschapsmensen (van zeer gevarieerde samenstelling) bijeen, van wie een aantal na afloop van de oorlog de toen ontwikkelde technieken op bedrijfsproblemen ging toepassen. En niet zonder succes.

De bedoeling van dit boek is een inzicht te geven in hetgeen econometrie en operationele research presteren. Ter geruststelling zij meteen al opgemerkt, dat dit verslag niet bijzonder technisch van aard is. Dat wordt trouwens al direct duidelijk voor wie dit boek doorbladert; de frequentie van formules per pagina is bepaald aan de geringe kant. Helemaal zonder wiskunde gaat het natuurlijk niet, maar we verlangen niet meer dan dat de lezer weet dat symbolen als x en y voor variabelen staan, dus voor dingen die kunnen variëren; dat a en b constanten zijn, dus dingen die niet variëren (bijv. 3 of 0 of —5); dat $y = a + bx$ een lineaire vergelijking is of een vergelijking van de eerste graad; dat $y = a + bx + cx^2$ een vergelijking van de tweede graad is; dat een kwadraat niet negatief kan zijn, dus alleen maar nul of positief [1]; dat hij nog weet wat een vierkantswortel is; en dat een vergelijking van de eerste graad kan worden weergegeven door een rechte lijn en een van de tweede graad door een kromme. Mocht de lezer zich ook op deze terreinen onzeker voelen, dan adviseren wij hem met Hoofdstuk 1 te beginnen, waar een aantal van deze zaken op elementaire wijze wordt uitgelegd. (Het advies om met Hoofdstuk 1 te beginnen doet wellicht ietwat vreemd aan,

1. De mathematisch geavanceerde lezer zij geruststellend verzekerd, dat we ons in dit boek geen ondeugende uitstapjes buiten het reële domein veroorloven.

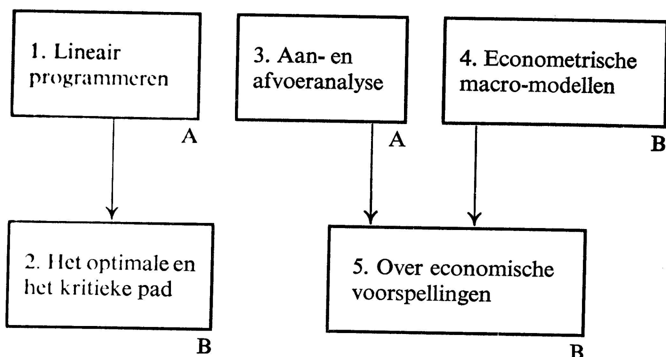
maar spoedig zal blijken, dat er tal van manieren zijn om dit boek te lezen.)

Het zal duidelijk zijn, dat men niet moet verwachten met behulp van dit boek een expert te worden. Wil men dat, dan is een academische studie op zijn plaats, waar men successievelijk wordt ingewijd in de economische problemen, in de wiskundige en statistische technieken die geschikt zijn om deze problemen tot oplossing te brengen, en in de oplossingprocedures zelf. Dat kost nu eenmaal tijd; bijv. een jaar of zes in de kwantitatief-economische studierichting van de Nederlandsche Economische Hoogeschool te Rotterdam. In dit boek stellen we ons ten doel een algemeen-begrijpelijk overzicht te geven, niet alleen met beperkte wiskundige middelen maar ook zonder een veronderstelde economische voorkennis, van een aantal belangrijke onderwerpen op het terrein van econometrie en operationele research. Dit heeft plaats in een dertiental korte hoofdstukken, die op blz. 12 in schematische opstelling met de titels zijn aangeduid. De pijlen hebben betrekking op de volgorde waarin ze gelezen dienen te worden; daarover hieronder meer. De letters rechtsonder geven de graad van moeilijkheid aan:

A. Eenvoudig
B. Iets minder eenvoudig
C. Nog iets minder eenvoudig.

Om misverstand te vermijden: deze driedeling heeft niet betrekking op de gebruikte wiskundige middelen. In Hoofdstuk 2, dus een B-hoofdstuk, gaat het erom dat het betoog als geheel zich over een vrij groot aantal pagina's uitstrekt, zodat enige inspanning (maar toch niet zo bijzonder veel!) vereist is om de grote lijn vast te houden. In Hoofdstuk 12, dus een C-hoofdstuk, is de gedachtengang gedeeltelijk nogal abstract. Enzovoorts.

Nu dan de pijlen. Wanneer er zo'n pijl gaat van het ene hoofdstuk naar het andere, betekent dit dat het ene hoofdstuk gelezen moet worden wil het andere begrijpelijk zijn. Dus Hoofdstuk 1 vóór Hoofdstuk 2; evenzo 3 en 4 vóór 5, maar 3 en 4 kunnen onmiddellijk gelezen worden evenals 6, enz. Dit betekent niet, dat in Hoofdstuk 6 in het geheel niet wordt terugverwezen naar eerdere hoofdstukken. Integendeel, dat komt wel voor, maar dergelijke terugverwijzingen dienen dan



6. Over onzekerheid en waarschijnlijkheid
B

7. Het strategie-begrip
A

9. Wachten
B

12. De statistische specificatie van economische relaties
C

8. Speltheorie
A

10. Simulatie en beleidsspelen
A

13. Om de gulden van de consument
B

11. Over productie- en voorraadbeslissingen
C

om het eerder behandelde in het kader van Hoofdstuk 6 te plaatsen en men kan daar zonder bezwaar over heen lezen als dat eerder behandelde niet tevoren gelezen is. Het schema laat zien, dat geen enkel hoofdstuk meer dan twee hoofdstukken 'voorstudie' vereist.

Twee pijlen zijn gestippeld getekend. Dit betekent zoveel als 'aanbevolen maar niet vereist'. Zo wordt er in Hoofdstuk 10 als toepassing van de simulatietechniek een wachttijdprobleem behandeld; het is nuttig Hoofdstuk 9 tevoren te lezen, maar aan de andere kant is het probleem zo doorzichtig, dat het niet werkelijk noodzakelijk is dit te doen. Evenzo, wat betreft het consumentengedrag in Hoofdstuk 13, wie bereid is over de standaardfouten (een maatstaf voor de betrouwbaarheid van de uitkomsten) heen te lezen kan het zonder Hoofdstuk 12 stellen. Maar hij mist natuurlijk wel iets.

Het boek wordt afgesloten met een Epiloog, een tableau de la troupe, waarin de diverse onderwerpen van de dertien hoofdstukken nogmaals aan de orde komen. Wie een globale indruk wil hebben van wat er in dit boek zo al ter sprake komt, kan dat op eenvoudige wijze doen door een blik te slaan op de eerste vier alinea's van de Epiloog.

1. LINEAIR PROGRAMMEREN

1. DE STRUCTUUR VAN EEN LINEAIR PROGRAMMERINGSPROBLEEM

Om de smaak te pakken te krijgen beschouwen we het volgende beroemd geworden dieetprobleem. Met menige huisvrouw vroeg G. J. Stigler zich omstreeks 1945 af wat het goedkoopste voedingspakket zou zijn nodig om een persoon een jaar lang in leven en goede gezondheid te houden. Uiteraard diende het menu aan een aantal voedingseisen te voldoen: het moest voldoende vitaminen bevatten; voldoende vetten en eiwitten; een zeker minimum aan calorieën en zo meer. Er waren 9 dergelijke voorwaarden. Het menu moest worden gekozen uit 77 verschillende voedingsmiddelen. Van elk van deze voedingsmiddelen was de prijs bekend. Bovendien was bekend hoeveel eiwit, vet, vitamine e.d. ieder product per eenheid bevat. Berekeningen op basis van deze gegevens leerden, dat het goedkoopste menu dat aan al die *negen* eisen voldeed bestond uit *negen* gerechten; deze overeenkomst in de getallen is, naar wij verder zullen zien, geen louter toeval. Het menu bevatte slechts verschillende hoeveelheden spinazie, pindakaas, reuzel, aardappelen, boerenkool, mais, tarwebloem, varkenslever en melkpoeder. Inderdaad niet zeer smakelijk, en waarschijnlijk wat moeilijk verteerbaar. Door de vraag wat strenger te formuleren en rekening te houden met smaak, verteerbaarheid en afwisseling werd een veel aantrekkelijker resultaat verkregen, waarvan de kosten echter bijna drie keer zo hoog waren. De oplossing van een variant van dit probleem – een menu samen te stellen met zo min mogelijk calorieën dat aan alle voedingstechnische eisen voldoet – kan men wekelijks vinden in de daartoe bestemde rubrieken der damesbladen.

De *structuur* van het vraagstuk is als volgt. De onbekenden zijn de hoeveelheden van de verschillende voedingsmiddelen die wij moeten kopen. Zoals gebruikelijk in de wiskunde geven wij deze onbekenden aan met een symbool, bijv. x; in dit geval, omdat er 77 verschillende voedingsmiddelen zijn, x_1, x_2,... tot en met x_{77}. Als het eerste goed levertraan is, het tweede radijs,... en het laatste rode bieten, dan betekent x_1 het aantal liters levertraan, x_2 het aantal bosjes radijs, ... en x_{77} het

aantal kilo's rode bieten dat wij kopen. Kortheidshalve zullen wij al deze hoeveelheden aangeven met x_i, waarbij de index i dus de waarden 1, 2, ..., 77 kan aannemen afhankelijk van de vraag welke van de 77 voedingsmiddelen het betreft.

Wij gaan nu de structuur van het probleem in wiskundige vorm gieten en beginnen met de doelstelling. Deze komt neer op het samenstellen van een menu met zo gering mogelijke kosten; het is dus onze taak een uitdrukking voor deze kosten te vinden. Daartoe geven we de prijzen van de verschillende voedingsmiddelen aan met p_i. Dus betekent p_1 de prijs in guldens van een liter levertraan, p_2 de prijs in guldens van een bosje radijs, ... en p_{77} de prijs van een kilo rode bieten. Gemakkelijk zien we dan: als x_1 liter levertraan wordt gekocht, wordt daarvoor p_1x_1 gulden uitgegeven; als x_2 bosjes radijs worden gekocht, dan kost dat p_2x_2 gulden. Beide tezamen kosten dan $p_1x_1 + p_2x_2$ (gulden). Op deze wijze zien wij dat de *totale* kosten van een voedingspakket bestaande uit x_1 liter levertraan, x_2 bosjes radijs, ... en x_{77} kilo rode bieten gelijk zijn aan

$$K = p_1x_1 + p_2x_2 + \ldots + p_{77}x_{77}.$$

Dit wordt de *doelstellingsfunctie* genoemd. De waarde K moet worden geminimeerd. De doelstellingsfunctie is *lineair* in de onbekenden x_i, d.w.z. van de vorm 'zoveel maal x_1 plus zoveel maal x_2 plus ...'.[1]

Hoe moeten we ons de oplossing van dit minimumprobleem

1. Een lineaire functie (zonder 'constante term') van een aantal variabelen x_1, x_2, ... heeft de volgende eenvoudige eigenschap: als we al die variabelen x_i tweemaal zo groot nemen, dus vervangen door $2x_i$, dan wordt de functiewaarde zelf ook tweemaal zo groot; als we x_i vervangen door $7x_i$, wordt ook de functiewaarde met 7 vermenigvuldigd; enz. In ons geval:

$$\begin{aligned} &p_1(2x_1) + p_2(2x_2) + \ldots + p_{77}(2x_{77}) \\ &= 2p_1x_1 + 2p_2x_2 + \ldots + 2p_{77}x_{77} \\ &= 2(p_1x_1 + p_2x_2 + \ldots + p_{77}x_{77}) = 2K. \end{aligned}$$

Dit staat in tegenstelling tot bijv. kwadratische functies. Een voorbeeld daarvan: $3x^2$ is een kwadratische functie van x. Vervangen we x door $2x$, dan wordt de functiewaarde:

$$3(2x)^2 = 4 \times 3x^2$$

dus niet 2 maar 4 keer zo groot.

voorstellen? Wordt het bereikt door alle x_i nul te maken? Zeker, de kosten zijn dan bijzonder laag, nl. ook nul. Maar aan de voorwaarden inzake vitaminen, calorieën e.d. wordt dan niet voldaan.

Deze voorwaarden, welke als *nevenvoorwaarden* worden aangeduid, kunnen we als volgt weergeven. Als voorbeeld beschouwen we de calorieën-restrictie. Op grond van medische overwegingen heeft een volwassene tenminste 800.000 calorieën per jaar nodig. Het is gegeven dat 1 liter levertraan c_1 calorieën bevat, een bosje radijs c_2 calorieën, . . . en 1 kilo rode bieten c_{77} calorieën. Als men x_1 liter levertraan, x_2 bosjes radijs, . . . en x_{77} kilo rode bieten koopt heeft men in totaal

$$c_1x_1 + c_2x_2 + \ldots + c_{77}x_{77}$$

calorieën. Dit totaal nu dient *tenminste* 800.000 te bedragen, zodat wij schrijven

$$c_1x_1 + c_2x_2 + \ldots + c_{77}x_{77} \geqq 800.000,$$

waarbij $\geqq$ staat voor 'tenminste gelijk aan' of ook 'groter dan of gelijk aan'. (Het teken $>$ wordt algemeen gebruikt voor 'groter dan'.) Soortgelijke restricties gelden voor vitaminen, vetten, eiwitten; in totaal zijn er in het onderhavige probleem 9 nevenvoorwaarden. Merk op, dat het linkerlid van een dergelijke restrictie lineair is in de x_i, precies als het geval is bij de doelstellingsfunctie.

We zijn er nu bijna, nog niet helemaal. Er moet nl. nog expliciet gesteld worden, dat de x_i niet negatief mogen zijn. We kunnen immers geen negatief aantal liters levertraan kopen, geen negatief aantal bosjes radijs! Dit lijkt vanzelfsprekend en is het ook; maar men moet goed beseffen dat het wiskundige apparaat er zich niets van aantrekt of een restrictie vanzelfsprekend is of niet. Dus leggen we op:

$$x_i \geqq 0 \quad \text{voor} \quad i = 1, 2, \ldots, 77,$$

in woorden: de gekochte hoeveelheden van alle 77 voedingsmiddelen moeten hetzij nul zijn, hetzij positief. Deze voorwaarden staan bekend als de *niet-negativiteitsrestricties.*

Hiermee hebben wij het terrein van de lineaire programmering afgebakend. Het gaat er steeds om de waarden van een aantal variabelen vast te stellen. In ons voorbeeld: de 77 hoeveelheden van de diverse voedingsmiddelen. Deze variabelen

worden beheerst, d.w.z. zij kunnen worden vastgesteld, door de persoon voor wie het probleem een probleem is (in dit geval de huisvrouw die de inkopen doet). Deze persoon moet dus beslissen omtrent de door die variabelen aan te nemen waarden, vandaar dat zij *beslissingsvariabelen* worden genoemd. Volgens welk criterium heeft de beslissing plaats? Zodanig, dat een *lineaire doelstellingsfunctie* in deze variabelen wordt geminimeerd. In ons voorbeeld de totale kosten van het menu. De minimering heeft echter niet 'onvoorwaardelijk' plaats, want de oplossing dient te voldoen (i) aan de niet-negativiteitsrestricties voor alle beslissingsvariabelen en (ii) aan een aantal nevenvoorwaarden – inzake eiwit, vitaminen e.d. in het hier beschouwde voorbeeld. Die restricties hebben de vorm van 'ongelijkheden', d.w.z. een zekere uitdrukking dient tenminste gelijk te zijn aan een bepaald bedrag (zie onze calorieën-restrictie). Het kan ook zijn 'ten hoogste gelijk', hetgeen bijv. het geval geweest zou zijn wanneer we ook een restrictie hadden ingevoerd van het type: het is medisch ongewenst wanneer de persoon *meer* dan 2 miljoen calorieën per jaar tot zich neemt. Tenslotte geldt bij lineair programmeren, dat de ongelijkheden lineair in de beslissingsvariabelen zijn.

Soms komt het voor (wij zullen hieronder een voorbeeld zien), dat er niet geminimeerd maar gemaximeerd wordt. Het is van belang in te zien, dat dit wiskundig nauwelijks enig verschil uitmaakt. Neem daartoe de kosten van het menu,

$$K = p_1x_1 + p_2x_2 + \dots + p_{77}x_{77},$$

die dus geminimeerd dienen te worden. Vermenigvuldig beide zijden met —1:

$$-K = -p_1x_1 - p_2x_2 - \dots - p_{77}x_{77}.$$

Deze uitdrukking, dus minus de kosten, moet worden gemaximeerd! Neem bijv. een menu van 30 gulden per week voor een gezin; neem daarnaast het optimale menu, dus het goedkoopste (dat evenals het eerste aan alle eisen voldoet) en laat het bijv. 28 gulden kosten. Dan zijn minus de kosten —30 resp. —28 gulden en het optimale menu heeft dus een (algebraïsch) groter negatief kostenbedrag; immers —28 is groter dan —30. Waaruit volgt, dat we elk minimeringsprobleem tot een maximeringsprobleem kunnen omvormen door de doelstellingsfunctie van een minteken te voorzien. Zo kunnen we natuurlijk ook een maximumprobleem tot een minimumprobleem herleiden.

Een probleem van kleinere afmetingen (met maar 2 beslissingsvariabelen i.p.v. 77) is nuttig om de oplossingstechniek nader te bestuderen. Stellen we ons daartoe een radiofabrikant voor die slechts twee soorten radio's produceert; een eenvoudige uitvoering – die desondanks een 'perfecte geluidsweergave' garandeert – en een luxe uitvoering, met meer knoppen en iets groter, – 'het laatste woord in de radiotechniek' –. De radio's worden verkocht met een winstmarge van 20, resp. 30 gulden per stuk. De beslissingsvariabelen zijn het *aantal* per dag te produceren eenheden van de goedkope eenvoudige radio, r_1, en het *aantal* per dag te produceren eenheden van de duurdere luxe radio, r_2. Als de radioproducent besluit r_1 eenvoudige radio's per dag te produceren geeft hem dit een winst van $20r_1$ gulden per dag; r_2 luxe radio's leveren hem $30r_2$ gulden. De fabrikant wenst dus

$$20r_1 + 30r_2$$

te maximeren; immers $20r_1 + 30r_2$ stelt zijn *totale winst* voor. Uiteraard mogen r_1 en r_2 niet negatief zijn. Men kan nu eenmaal geen negatief aantal radio's produceren. In eerste instantie is men geneigd te denken dat men in zulk geval uitsluitend luxe radio's moet produceren. Er zijn echter *capaciteitsgrenzen* gesteld aan de productie; deze geven aanleiding tot nevenvoorwaarden.

Stel U daartoe voor, dat de productie plaats vindt aan twee lopende banden; één voor de eenvoudige radio, en één voor de luxe radio. De capaciteit van deze banden is beperkt. Men kan per dag maximaal 7 eenvoudige radio's produceren:

$$r_1 \leqq 7,$$

waarbij $\leqq$ staat voor 'ten hoogste gelijk aan' of ook 'kleiner dan of gelijk aan'. (Het teken $<$ staat voor 'kleiner dan'.) Evenzo veronderstellen we, dat per dag maximaal 5 luxe radio's kunnen worden gemaakt:

$$r_2 \leqq 5.$$

Bovendien is er een nevenvoorwaarde omdat de beschikbare hoeveelheid arbeid beperkt is. Ware dit niet zo, dan zou men de banden beide op volle toeren kunnen laten draaien, en derhalve per dag 7 eenvoudige en 5 luxe radio's produceren. Dit kan echter niet. Er werken slechts 12 werknemers in de fabriek, zodat er per dag niet meer dan 12 'mandagen' ter be-

schikking staan. Een eenvoudige radio vereist 1 mandag arbeid, een luxe radio verlangt 2 mandagen arbeid in de loop van zijn productie. Maken we r_1 eenvoudige radio's per dag, dan vereist dit dus r_1 mandagen; maken we r_2 luxe radio's per dag, dan zijn daartoe $2r_2$ mandagen noodzakelijk. Het totale programma vereist kennelijk $r_1 + 2r_2$ mandagen, zodat we geconfronteerd worden met de volgende nevenvoorwaarde:

$$r_1 + 2r_2 \leqq 12.$$

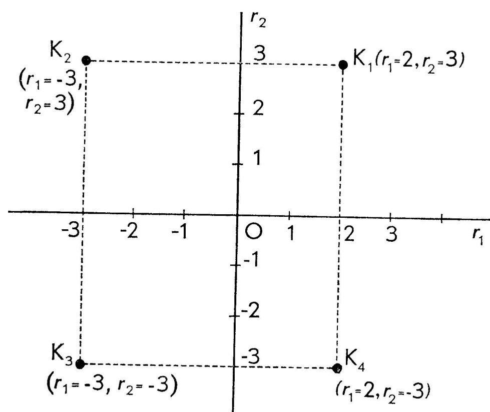
Merk op, dat dit het op volle toeren laten lopen van de twee banden onmogelijk maakt. Dan immers zou gelden $r_1 = 7$ en $r_2 = 5$, zodat $r_1 + 2r_2 = 17$; en die 17 mandagen zijn eenvoudig niet beschikbaar. Wèl kan men bijv. 4 eenvoudige en 4 luxe radio's maken (dan zijn alle employé's vol bezet, want $4 + 2 \times 4 = 12$), hetgeen aanleiding zou geven tot een winst van $4 \times 20 + 4 \times 30 = 200$ gulden per dag. Wat weer niet kan is alle employé's aan de productie van luxe radio's te werk te stellen. Dit zou betekenen $r_1 = 0$ (er worden geen eenvoudige radio's gemaakt) en $r_2 = 6$; gegeven de 2 mandagen noodzakelijk voor een luxe radio gaan alle 12 beschikbare mandagen daar precies aan op. Maar het is onmogelijk per dag 6 luxe radio's te maken i.v.m. de capaciteitsrestrictie van de desbetreffende lopende band, $r_2 \leqq 5$.

Het is eenvoudig in te zien, dat dit productievoorbeeld valt onder de lineaire programmering. We hebben 2 beslissingsvariabelen, r_1 en r_2, die beide door de radiofabrikant naar believen kunnen worden vastgesteld – althans binnen zekere grenzen. Deze grenzen worden enerzijds door de niet-negativiteitsrestricties bepaald ($r_1 \geqq 0$, $r_2 \geqq 0$), anderzijds door de nevenvoorwaarden. Van deze laatste hebben er twee betrekking op de capaciteitsgrenzen van de lopende banden ($r_1 \leqq 7$, $r_2 \leqq 5$), de derde op de beschikbare mankracht; maar voor alle drie geldt, dat het linkerlid lineair is in r_1 en r_2. Dit laatste geldt ook voor de doelstellingsfunctie $20r_1 + 30r_2$ die wij wensen te maximeren. Hiermee is dus inderdaad aangetoond dat we met een geval van lineair programmeren te maken hebben.

3. EEN GRAFISCHE AFBEELDING

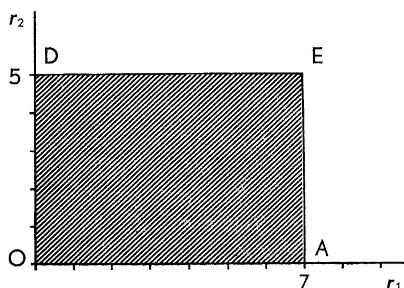
Wanneer er slechts twee variabelen zijn, zoals in ons radiovoorbeeld, dan loont het bepaald de moeite om het probleem

in grafische vorm weer te geven. We krijgen dan nl. een veel duidelijker beeld van wat er precies aan de hand is. Laten we dan een assenkruis opstellen met twee loodrecht op elkaar staande assen; langs de horizontale as meten we r_1 en langs de verticale as r_2.



Neem dan het punt K_1. Aan de hand van de stippellijnen, die vanuit dit punt evenwijdig aan beide assen zijn getrokken, zien we dat K_1 ligt op een afstand 2 van de verticale as en op een afstand 3 van de horizontale as. Dit betekent, dat K_1 de grafische voorstelling is van het productieprogramma $r_1 = 2$, $r_2 = 3$. Elk punt in het vlak is een afbeelding van een productieprogramma; omgekeerd wordt elk programma door een punt afgebeeld. Maar sommige punten stellen heel vreemde productieprogramma's voor. Neem bijv. K_2; daar geldt $r_1 = -3$ en $r_2 = 3$. Kennelijk wordt in K_2 een negatief aantal eenvoudige radio's gemaakt, hetgeen door gezond verstand (wiskundig geformaliseerd door de niet-negativiteitsrestrictie) is uitgesloten. Dit geldt voor *alle* punten, die links van de verticale as liggen; de voorwaarde $r_1 \geqq 0$ vereist dat we hetzij op de verticale as hetzij rechts daarvan zitten. Op geheel overeenkomstige wijze impliceert de voorwaarde $r_2 \geqq 0$, dat we hetzij op de horizontale as hetzij daarboven dienen te zitten. (Gemakkelijk ziet men, dat de punten K_3 en K_4 beide deze voorwaarde schenden; hun r_2 is negatief en zij liggen onder de horizontale as.)

De twee niet-negativiteitsrestricties tezamen impliceren dus het volgende voorschrift: zoek de oplossing *op of boven* de horizontale as en *op of rechts* van de verticale as. Daar is het vrije jachtterrein. Maar ook deze vrijheid is beperkt. We hebben nl. de capaciteitsrestricties van de twee lopende banden, $r_1 \leqq 7$ en $r_2 \leqq 5$. Het punt K_1 van onze figuur voldoet hieraan; daarvoor geldt immers $r_1 = 2$, $r_2 = 3$. Maar zouden we een punt nemen dat verder naar rechts ligt dan een afstand 7 van de verticale as, of verder naar boven dan een afstand 5 van de horizontale, dan is het mis; het afgebeelde productieprogramma kan dan niet door de lopende banden gerealiseerd worden. Dit leidt tot de volgende figuur:



Kennelijk beperkt de vrijheid van keuze zich tot de gearceerde rechthoek. (Gemakshalve zijn de negatieve stukken van de assen weggelaten; de niet-negativiteitsrestricties maken dat die voor ons doel niet relevant zijn.)

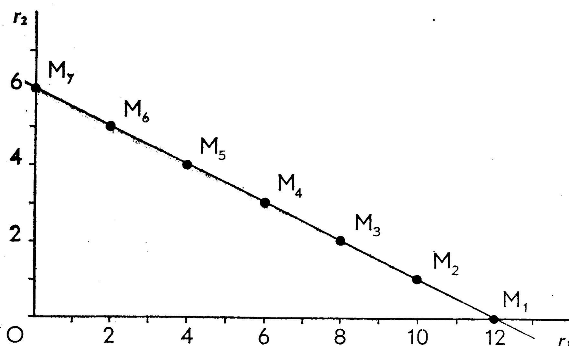
Voor hetgeen volgt is het van belang nog even nauwkeurig aan te geven hoe wij aan de gearceerde rechthoek zijn gekomen. Die rechthoek wordt bepaald door zijn vier zijden; en die zijden corresponderen op hun beurt met de vier restricties $r_1 \geqq 0$, $r_1 \leqq 7$; $r_2 \geqq 0$, $r_2 \leqq 5$. Om de voorwaarde $r_1 \leqq 7$ (r_1 *is kleiner dan of gelijk aan* 7) af te beelden tekenen wij eerst de lijn waarop geldt $r_1 = 7$ (r_1 *is gelijk aan* 7). Dit is de lijn *AE* in de figuur. Voor ieder punt op *AE* geldt inderdaad $r_1 = 7$. Zouden we deze lijn naar boven en naar onder onbepaald ver doortrekken, dan wordt daardoor het hele vlak in twee helften verdeeld; de éne kant is goed, toelaatbaar wat die nevenvoorwaarde betreft; de andere is fout, niet toegelaten. Want aan die andere kant geldt $r_1 > 7$ (r_1 *is groter dan* 7). Iets analoogs

geldt voor willekeurige lineaire ongelijkheden. Teken *eerst* de lijn waar de corresponderende gelijkheid precies opgaat. Deze lijn verdeelt het vlak in twee delen, waarvan er één deel bestaat uit alle toelaatbare punten, en het andere uit uitsluitend ontoelaatbare punten.

Deze procedure nu zullen wij ook toepassen op de nevenvoorwaarde inzake het aantal beschikbare mandagen, dus $r_1 + 2r_2 \leqq 12$. Ook hier houden we ons eerst bezig met het geval waar alle employés volledig bezet zijn, dus $r_1 + 2r_2 = 12$. Het is eenvoudig een aantal combinaties van waarden r_1 en r_2 te vinden zodanig dat $r_1 + 2r_2$ gelijk is aan 12. Bijvoorbeeld:

$$\begin{array}{llll}
r_1 = 12, & r_2 = 0 & (12 + 2 \times 0 = 12); & M_1 \\
r_1 = 10, & r_2 = 1 & (10 + 2 \times 1 = 12); & M_2 \\
r_1 = 8, & r_2 = 2 & (8 + 2 \times 2 = 12); & M_3 \\
r_1 = 6, & r_2 = 3 & (6 + 2 \times 3 = 12); & M_4 \\
r_1 = 4, & r_2 = 4 & (4 + 2 \times 4 = 12); & M_5 \\
r_1 = 2, & r_2 = 5 & (2 + 2 \times 5 = 12); & M_6 \\
r_1 = 0, & r_2 = 6 & (0 + 2 \times 6 = 12); & M_7
\end{array}$$

Wij hebben aldus 7 verschillende productieprogramma's, die door 7 punten $M_1, M_2, \ldots, M_7$ worden afgebeeld. Dit is gedaan in de volgende figuur:



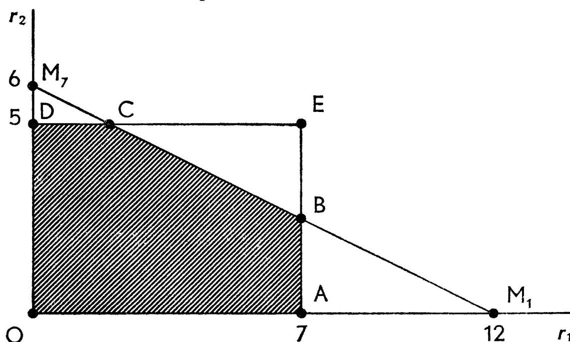
Wij zien, dat *ook in dit geval* alle punten op een rechte lijn liggen. Dat is een algemene eigenschap van lineaire vergelijkingen in twee variabelen: *hun grafische voorstelling is een rechte lijn.* Dus óók bijvoorbeeld $6r_1 + r_2 = 18$, of $27r_1 + 3r_2 = 81$;

zij worden alle afgebeeld als een rechte lijn (in het platte vlak, met assen waarlangs r_1 en r_2 gemeten worden).

Omdat een rechte lijn door *twee* punten volledig bepaald is, is de constructie ervan al bijzonder eenvoudig. Men bepale slechts (i) hoe groot r_1 is als r_2 nul is; en (ii) hoe groot r_2 is als r_1 nul is. Men heeft dan twee punten, en dus de lijn. In dit voorbeeld hadden we kunnen volstaan met het bepalen van M_1 ($r_1 = 12$ als $r_2 = 0$) en M_7 ($r_2 = 6$ als $r_1 = 0$).

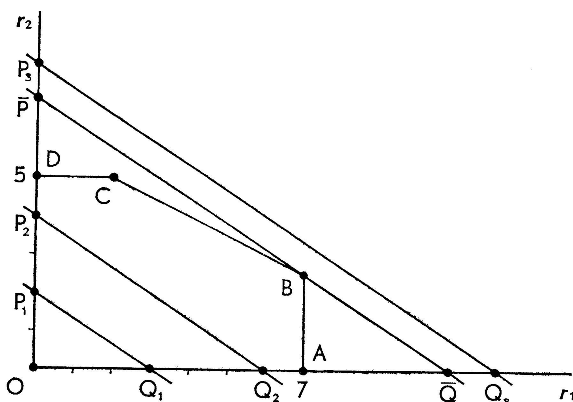
Wij hebben nu de lijn M_1M_7 waar geldt $r_1 + 2r_2 = 12$. Deze lijn verdeelt het vlak in twee helften. Voor álle punten in de éne helft geldt $r_1 + 2r_2 < 12$ – in déze helft zijn we geïnteresseerd. Voor álle punten in de andere helft geldt $r_1 + 2r_2 > 12$ – voor deze helft hebben we geen belangstelling. Wat nog resteert is het vaststellen van de 'goede' helft. Dit kan het eenvoudigst door na te gaan of de oorsprong O (het snijpunt van de twee assen), waar $r_1 = 0$ en $r_2 = 0$, in de *goede* of in de *foute* helft ligt. Ligt O in de goede helft (zoals hier, want $0 + 2 \times 0 = 0 < 12$), dan is de helft die O bevat de helft die we moeten hebben. Anders is het de andere helft waarin we zijn geïnteresseerd. Voor ons probleem moeten we dus zitten *onder* de lijn door M_1 en M_7.

We hebben onze bouwstenen gereed en het wordt tijd ze te verenigen. Wil voldaan zijn aan de niet-negativiteitsrestricties en aan de capaciteitsrestricties van de lopende banden, dan dienen we de gearceerde rechthoek in de figuur op blz. 21 niet te verlaten. Wil voldaan zijn aan de nevenvoorwaarde inzake het beschikbare aantal mandagen, dan dienen we onder de rechte door M_1 en M_7 te zitten. Aan beide voorwaarden dient voldaan te zijn en dus krijgen we de volgende figuur:



Het ‘toelaatbare gebied’ is nu ingekrompen tot een vijfhoek. Linksonder hebben we de oorsprong *O*. Rechtsonder hebben we *A;* zouden we nog verder naar rechts gaan, dan wordt de capaciteitsgrens van de lopende band voor eenvoudige radio’s overschreden. Daarboven hebben we *B;* zouden we verder omhoog gaan, dan gebruiken we meer mandagen dan ter beschikking staan. Linksboven *B* hebben we *C;* zouden we de tocht naar linksboven verder vervolgen, dan overschrijden we de capaciteitsgrens van de lopende band voor luxe radio’s. Links van *C* hebben we *D*; nog verder links wordt het aantal eenvoudige radio’s negatief.

Hiermee zijn de nevenvoorwaarden en de niet-negativiteitsrestricties afgedaan. Wij gaan nu over tot de doelstellingsfunctie, dus de winst $20r_1 + 30r_2$ die wij wensen te maximeren. Daartoe bezien we eerst het speciale geval van een winst van 60 gulden per dag. Met andere woorden, wij vragen ons af voor welke combinaties van r_1 en r_2 geldt $20r_1 + 30r_2 = 60$. Dit is weer een lineaire vergelijking (van hetzelfde type als $r_1 + 2r_2 = 12$), die wordt weergegeven door een rechte lijn. De betreffende rechte gaat door $r_1 = 3$ als r_2 gelijk nul is; en door $r_2 = 2$ als r_1 nul is. Zo zijn P_1 en Q_1, en daarmee ook de lijn P_1Q_1, in de nu volgende figuur bepaald. *Voor ieder punt op* P_1Q_1 *bedraagt de totale winst* 60.



Merk op, dat de lijn P_1Q_1 gedeeltelijk in het toelaatbare gebied ligt, gedeeltelijk daarbuiten; alleen in het eerste geval heb-

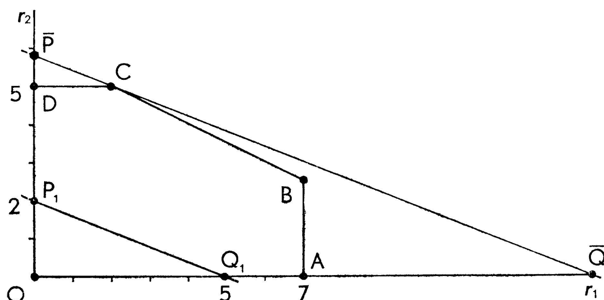
ben we te doen met een realiseerbaar productieprogramma.

Vervolgens houden we ons bezig met een winst van 120 gulden per dag. Deze wordt gerealiseerd door bijv. $r_1 = 6$, $r_2 = 0$ en $r_1 = 0$, $r_2 = 4$; ook nu liggen al dergelijke punten op een rechte lijn, te weten P_2Q_2 in dezelfde figuur. Deze rechte ligt verder van de oorsprong verwijderd dan P_1Q_1, hetgeen niets anders is dan de meetkundige betekenis van het feit dat de winst groter is: 120 gulden i.p.v. 60. (In de oorsprong zèlf is de winst nul!) Kennelijk gaat het erom, op een winstlijn te zitten die *zo ver mogelijk* van de oorsprong verwijderd is; dan immers wordt de winst gemaximeerd. Maar dat kunnen we niet onbeperkt doen. Neem bijv. de lijn P_3Q_3; de winst is daar 240 gulden, maar de lijn bevat geen enkel punt in het toelaatbare gebied! Er is dus geen enkel programma met een winst van 240 dat kan worden geproduceerd. Blijkbaar dienen we *die* winstlijn op te sporen, die enerzijds zo ver mogelijk van de oorsprong ligt maar anderzijds nog net een punt in het toelaatbare gebied heeft. Dit is de lijn $\bar{P}\bar{Q}$ en het punt in kwestie is B, dat derhalve de oplossing van het lineaire programmeringsprobleem vormt. De winst in B is 215 gulden per dag; immers $7 \times 20 + 2^1/_2 \times 30 = 215$. Dit is dus de maximaal haalbare winst. Deze wordt bereikt door per dag 7 eenvoudige en $2^1/_2$ luxe radio's te maken (dus 5 luxe radio's per 2 dagen).

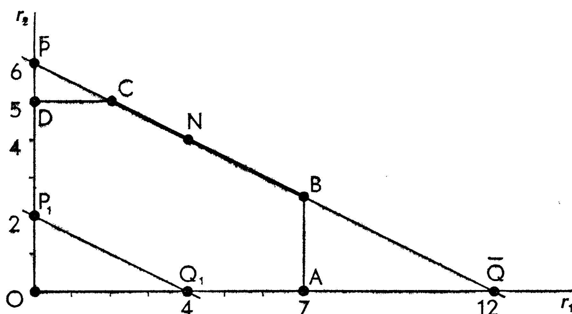
4. DE OPLOSSING BIJ VERSCHILLENDE WINSTMARGES

Ons probleem is aldus grafisch opgelost, hetgeen zo eenvoudig kon omdat er slechts twee beslissingsvariabelen (r_1 en r_2) waren. Maar niet alleen *dit* probleem is opgelost. Als de *winstmarges* zouden veranderen, zou de nieuwe oplossing óók direct bepaald kunnen worden. Bijv., als de winst op een eenvoudige radio 20 gulden blijft, doch die op een dure radio 50 gulden wordt, dan dient men $20r_1 + 50r_2$ te maximeren. De lijnen PQ gaan nu minder steil lopen en het gevolg is dat niet B maar C de oplossing wordt. Deze situatie is grafisch geïllustreerd in de eerste van de twee figuren op blz. 26. In de nieuwe oplossing (C) produceert men 2 eenvoudige en 5 luxe radio's en de winst is dan $2 \times 20 + 5 \times 50 = 290$ gulden per dag. Zou men in de nieuwe situatie de oude oplossing

(B) aanhouden, dan is de winst slechts $7 \times 20 + 2^1/_2 \times 50 =$ 265 gulden per dag. In de nieuwe oplossing worden meer luxe radio's en minder eenvoudige radio's geproduceerd dan in de oude. Dat ligt in de rede, want de winstgevendheid van luxe radio's is gestegen vergeleken met die van het eenvoudige type.



Een heel speciaal geval treedt op, wanneer de winstmarge op luxe radio's gelijk is aan 40 gulden. De te maximeren uitdrukking wordt dan $20r_1 + 40r_2$ en de winstlijnen lopen evenwijdig aan *BC:*



Indien men de winstlijn zo ver mogelijk van de oorsprong verschuift komt er een moment, waarop die lijn met BC samenvalt. Verdere verschuivingen hebben *geen enkel punt* meer met het toelaatbare gebied gemeen. *Elk* punt op het lijnstuk BC, inclusief B en C zelf, geven nu een gelijke winst. In B: $7 \times 20 + 2^1/_2 \times 40 = 240$; in C: $2 \times 20 + 5 \times 40 = 240$.

Als men 4 exemplaren van beide soorten radio's produceert: $4 \times 20 + 4 \times 40 = 240$. Dit is het punt N van de figuur, tussen B en C in. Blijkbaar is in het geval van een winstmarge van 40 gulden op de luxe radio's het optimale productieprogramma niet eenduidig bepaald. Er zijn talrijke oplossingen die elk voldoen aan de nevenvoorwaarden en de niet-negativiteitsrestricties en die alle gekenmerkt zijn door *dezelfde* maximale winst.

We hebben hiermee onze grafische exploraties voltooid en zijn nu in staat de balans op te maken. In alle onderzochte gevallen ligt de oplossing op de *grens* van het toelaatbare gebied, dus niet middenin. Dit is steeds het geval bij lineair programmeren. Bovendien ligt de oplossing meestal in een *hoekpunt* (B of C in onze voorbeelden). De oplossing kan ook een compleet lijnstuk zijn, zoals BC in het laatst onderzochte geval. Maar zelfs dan maakt een hoekpunt deel uit van de oplossing: in ons voorbeeld twee stuks, nl. B en C. Wij concluderen: wat de winstmarges ook mogen zijn, *de* of althans *een* oplossing ligt op een hoekpunt van het toelaatbare gebied.

5. SPELINGSVARIABELEN

Werkelijk interessante lineaire programmeringsproblemen zijn gekenmerkt door een groter aantal beslissingsvariabelen dan twee. In deze richting zullen wij nu gaan werken.

Beschouw dan een willekeurige nevenvoorwaarde, zegge

$$x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 \leqq 19,$$

die bijv. als volgt geïnterpreteerd kan worden. Stel ik heb slechts 19 gulden ter beschikking (het rechterlid). Daarmee kan ik

x_5 keer naar een restaurant (kosten f 5, inclusief)
x_4 keer naar een concert (kosten f 4)
x_3 keer naar een theater (kosten f 3)
x_2 keer naar een film (kosten f 2)
x_1 keer naar een café (kosten f 1).

Ik kan dus niet 1 keer naar de film ($x_2 = 1$), 2 keer naar een concert ($x_4 = 2$) en 2 keer naar een restaurant ($x_5 = 2$), want $2 \times 1 + 4 \times 2 + 5 \times 2 = 20$, dus meer dan de beschikbare 19. Wél echter kan ik alles eenmaal doen, want $1 + 2 + 3 + 4 + 5 = 15$, dus houd ik nog 4 gulden over.

We gaan nu een zesde grootheid invoeren, x_6, die betekent: het aantal guldens dat *niet* wordt uitgegeven. Onze nevenvoorwaarde kan dan worden geschreven als

$$(x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5) + x_6 = 19.$$

Daarbij staat de uitdrukking tussen haken voor het bedrag dat wèl, de x_6 voor het bedrag dat nièt wordt uitgegeven; samen zijn zij natuurlijk 19 gulden. Merk op, dat wij aldus een ongelijkheid ($\leqq$) hebben omgevormd tot een gelijkheid ($=$), dus tot een vergelijking. Anders gezegd, door invoering van een nieuwe mogelijkheid – 'geld *niet* uitgeven' – is het *steeds mogelijk* een ongelijkheid door een gelijkheid te vervangen; en de nieuwe formulering is *volkomen gelijkwaardig met de oorspronkelijke formulering.* De nieuw ingevoerde variabele wordt 'spelingsvariabele' genoemd. Kennelijk is een spelingsvariabele evenals een beslissingsvariabele nooit negatief; we hebben immers $x_6 = 0$ als de 19 gulden precies wordt uitgegeven, $x_6 > 0$ als er geld overblijft, maar nooit $x_6 < 0$, want dit zou betekenen dat meer dan 19 gulden wordt uitgegeven, hetgeen juist door de nevenvoorwaarde wordt uitgesloten.

Door invoering van spelingsvariabelen kan men *alle* nevenvoorwaarden als gelijkheden i.p.v. ongelijkheden schrijven. In ons voorbeeld van de radiofabrikant kan men bijv. expliciet invoeren de spelingsvariabele r_3 met de volgende betekenis: de capaciteit van band 1 (voor eenvoudige radio's) die *niet* wordt gebruikt. Men krijgt dan i.p.v.

$$r_1 \leqq 7$$

(in woorden: de capaciteit van band 1 is beperkt tot 7 eenvoudige radio's per dag) de nieuwe restrictie:

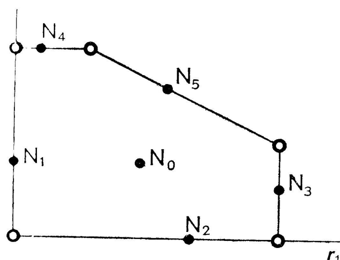
$$r_1 + r_3 = 7, \quad (r_3 \geqq 0).$$

Hetgeen wil zeggen dat de *gebruikte* capaciteit van band 1 plus de *niet gebruikte* capaciteit van band 1 is gelijk aan 7. Door evenzo in te voeren de spelingsvariabelen r_4 (niet gebruikte capaciteit van band 2) en r_5 (niet benutte mandagen) krijgt men de volgende nevenvoorwaarden:

$$\begin{aligned} r_1 \quad\quad + r_3 \quad\quad\quad\quad &= 7 \\ r_2 \quad\quad + r_4 \quad\quad &= 5 \\ r_1 + 2r_2 \quad\quad\quad\quad + r_5 &= 12. \end{aligned}$$

Dit is een stelsel van 3 gelijkheden in 5 onbekenden. Uiteraard zijn we alleen geïnteresseerd in oplossingen waarbij *géén* der variabelen negatief is. Al deze oplossingen corresponderen met een punt in de gearceerde vijfhoek van de figuur op blz. 23.

Omgekeerd, ieder punt in de gearceerde vijfhoek correspondeert met een niet-negatieve oplossing van het stelsel gelijkheden. Men merke nu op dat voor punten *binnen* de vijfhoek (dus niet op de rand) alle 5 onbekenden positief (dus niet nul) zijn. Een grafisch beeld is weer het eenvoudigst:



Voor N_0 geldt kennelijk, dat r_1 en r_2 beide positief zijn. Bovendien ligt N_0 links van de (verticale) capaciteitsrestrictie voor eenvoudige radio's en onder de (horizontale) capaciteitsrestrictie voor luxe radio's; in N_0 zijn *beide* lopende banden dus gekenmerkt door onvolledig gebruikte capaciteit, dus zijn r_3 en r_4 beide positief. Tenslotte ligt N_0 linksonder de neergaande rechte die de arbeidsrestrictie representeert; dus zijn er onbenutte mandagen, dus is ook r_5 positief. Maar neem nu N_1; daar geldt dat r_1 nul is (er worden geen eenvoudige radio's geproduceerd), terwijl r_2, r_3, r_4, r_5 alle positief zijn. Op overeenkomstige wijze geldt, dat in N_2, N_3, N_4, N_5 alle r's positief zijn met uitzondering van resp. r_2, r_3, r_4, r_5. Neemt men een van de vijf hoekpunten (in de figuur met cirkeltjes aangeduid), dan zijn *twee* van de r's nul en de overige drie positief. Bijvoorbeeld, in het hoekpunt rechtsonder is $r_2 = r_3 = 0$ en r_1, r_4, r_5 zijn positief.

Nu weten wij uit het voorgaande, dat de (of althans een) optimale oplossing van het lineaire programmeringsprobleem altijd met een hoekpunt correspondeert. We zien nu, dat in zo'n hoekpunt van de 5 variabelen er slechts 3 positief zijn en de andere 2 nul. Die 3 positieve variabelen zijn precies even talrijk als de nevenvoorwaarden. Dit nu geldt algemeen: *wanneer we m nevenvoorwaarden hebben in n beslissingsvariabelen* ($m = 3$, $n = 2$ bij de radio's), *dan hebben we $m + n$ beslissings- en spelingsvariabelen in totaal,* omdat elk van de m

nevenvoorwaarden aanleiding geeft tot één spelingsvariabele; *en dan geldt voor zo'n hoekpunt, dat m van die m + n variabelen positief zijn en de overige n nul.* Dus inderdaad evenveel positieve variabelen als er nevenvoorwaarden zijn.

Dit is geheel in overeenstemming met hetgeen in het begin van dit hoofdstuk over het goedkoopste menu is vermeld. Daar ging het om 77 hoeveelheden van mogelijke voedingsmiddelen, dus $n = 77$. Er werden 9 nevenvoorwaarden opgelegd, dus $m = 9$. Dus $m + n = 86$. Het optimale menu voldeed zonder speling aan alle nevenvoorwaarden en bestond uit 9 voedingsmiddelen. 'Zonder speling' betekent dat alle spelingsvariabelen nul zijn; van de 77 beslissingsvariabelen zijn er $77 - 9 = 68$ nul, en van alle 86 beslissings- en spelingsvariabelen in totaal zijn er precies $m = 9$ positief. Uiteraard behoeft het niet steeds zo te zijn dat alle positieve variabelen in de oplossing *beslissings*variabelen zijn. Er kunnen ook één of meer positieve spelingsvariabelen zijn, hetgeen in het dieetprobleem zou betekenen, dat het optimale menu dan gekenmerkt is door bijv. meer dan het minimale aantal calorieën. Dit kan het gevolg zijn van de overige vereisten (eiwit enz.). Dan zal het menu natuurlijk minder gerechten omvatten (minder positieve beslissingsvariabelen).

We hebben ons hierboven bepaald tot spelingsvariabelen in het geval van restricties van het type 'ten hoogste gelijk aan' ($\leqq$). Gemakkelijk ziet men in, dat 'tenminste gelijk aan' ($\geqq$) op analoge wijze kan worden gehanteerd, maar dan dient men de spelingsvariabele af te trekken. Bijvoorbeeld, de nevenvoorwaarde

$$7x_1 + 3x_2 + 2x_3 \geqq 4$$

is volkomen gelijkwaardig met

$$7x_1 + 3x_2 + 2x_3 - x_4 = 4, \quad x_4 \geqq 0.$$

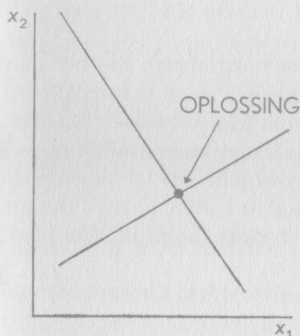
Zo'n spelingsvariabele (x_4) is dan 'het teveel van het een of ander' i.p.v. 'het niet gebruikte van het een of ander'.

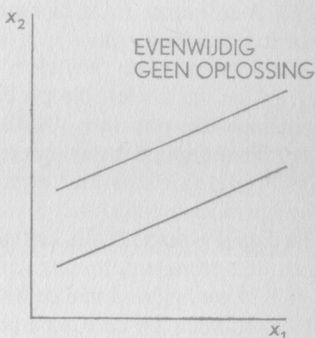
6. DE OPLOSSINGSFILOSOFIE VAN DE SIMPLEX-TECHNIEK

Nu zijn we dan zover, dat we ons op de eigenlijke oplossingsmethode kunnen gaan richten. D.w.z. de algebraïsch-numerieke methode, niet de grafische methode die ons zo behulpzaam was

bij het radio-voorbeeld; de grafische methode laat ons immers al gauw in de steek zodra het aantal variabelen gaat oplopen. Overigens zullen we de algebraïsch-numerieke methode niet in detail behandelen, want dat zou veel te ver voeren. We bepalen ons hier tot de meer fundamentele gedachtengang, die niet bijzonder ingewikkeld is. We hebben nl. $m + n$ beslissings- en spelingsvariabelen in totaal; we weten dat we naar hoekpunten moeten uitkijken; we weten dat in zo'n hoekpunt m variabelen positief zijn, de overige n nul; we hebben m lineaire vergelijkingen (de m nevenvoorwaarden als gelijkheden geschreven). Kennelijk moeten we dan n variabelen gelijk nul stellen en het stelsel van m vergelijkingen naar de overige m variabelen oplossen. Dan komen we automatisch op een hoekpunt uit! Immers, m lineaire vergelijkingen in m variabelen leiden, in de regel, tot precies één oplossing.

Om dit in te zien beschouwen wij als uitgangspunt twee lineaire vergelijkingen in twee onbekenden. Zo'n stelsel heeft gewoonlijk precies één oplossing. Dit is direct duidelijk als we de vergelijkingen tekenen. Wij krijgen dan 2 rechte lijnen en de eenduidige oplossing is hun snijpunt. Slechts als de twee lijnen toevalligerwijze *evenwijdig* zouden lopen, of (nog toevalliger), zouden *samenvallen* zijn er moeilijkheden.





Evenzo hebben 3 lineaire vergelijkingen in 3 onbekenden (gewoonlijk) precies één oplossing. Door het elimineren van een der onbekenden komt men n.l. bij het voorgaande geval terug. Op deze wijze vinden wij dat, in het algemeen, m lineaire vergelijkingen in m onbekenden *precies één oplossing toelaten*. Er

zijn weliswaar bijzondere gevallen (te vergelijken met het evenwijdig lopen of samenvallen van twee lijnen), doch deze zullen wij hier niet beschouwen.

De procedure is echter nog vrij bewerkelijk, omdat men de n variabelen die men van tevoren gelijk stelt aan nul op veel wijzen kan kiezen. In het radio-voorbeeld hadden wij een stelsel van 3 vergelijkingen in $2 + 3 = 5$ onbekenden. Wij moeten nu dus steeds *a priori* 2 variabelen gelijk stellen aan nul. Alle mogelijke wijzen om 2 variabelen uit 5 te kiezen – wiskundigen bewijzen eenvoudig dat er 10 mogelijkheden zijn[1] – kan men door met zorg en regelmaat te werk te gaan netjes opschrijven. Men kan dan 10 keer een stelsel van 3 vergelijkingen in 3 onbekenden oplossen. Soms leidt de oplossing tot negatieve waarden (dergelijke oplossingen vallen natuurlijk uit); de overige oplossingen worden ingevuld in de doelstellingsfunctie, waarna men kan nagaan welke tot de hoogste functiewaarde leidt.

Zoals gezegd is dit in de practijk nogal bewerkelijk. Er is echter een veel vlottere methode, de zgn. *Simplex-techniek*, die de volgende twee eigenschappen heeft:

(1) Men begint met een *niet-negatieve* oplossing die n variabelen gelijk nul heeft. Dit is gewoonlijk de oplossing waarbij slechts de spelingsvariabelen *niet* nul zijn. In ons voorbeeld: $r_1 = 0$, $r_2 = 0$, $r_3 = 7$, $r_4 = 5$, $r_5 = 12$. Anders gezegd, men begint in de oorsprong (waarin geen enkele radio wordt geproduceerd).

(2) Men gaat nu 'switchen' tussen variabelen die nul zijn (r_1, r_2) en variabelen die positief zijn (r_3, r_4, r_5). In elke stap ruilt men één positieve variabele met één nul-variabele. Men doet dit zodanig dat men zorgt dat de niet-nul-variabelen positief blijven (althans niet negatief worden) én dat de waarde van de doelstellingsfunctie van stap tot stap hoger (althans niet lager) wordt. Is dit laatste niet meer mogelijk, dan heeft men de optimale oplossing gevonden.

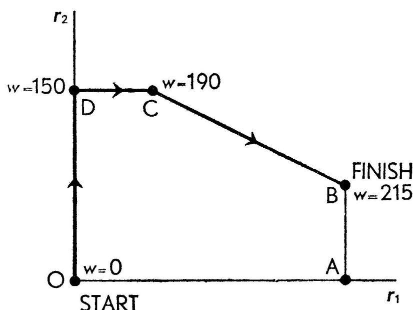
Dit is het beginsel van de Simplex-techniek. Gewoonlijk zijn er tussen m en $2m$ verwisselingen nodig voor men het eindpunt

1. Wellicht denkt men dat het er slechts 5 zijn op grond van het feit, dat de figuur op blz. 29 slechts 5 hoekpunten bevat. In werkelijkheid zijn er meer, omdat er bijv. ook een snijpunt is van de verticale as met de rechte die de arbeidsrestrictie representeert. (Van de hier genoemde 10 mogelijkheden vallen er in dit geval 2 uit, omdat er twee paar evenwijdige lijnen zijn.)

heeft bereikt. Als men in ons voorbeeld volgens de regels van de Simplex-kunst te werk gaat vindt men achtereenvolgens (w = winst = $20r_1 + 30r_2$);

I: $r_1 = 0, \quad r_2 = 0, \quad r_3 = 7, \quad r_4 = 5, \quad r_5 = 12$
(de oorsprong O, $w = 0$)
II: $r_1 = 0, \quad r_2 = 5, \quad r_3 = 7, \quad r_4 = 0, \quad r_5 = 2$
(het punt D, $w = 150$)
III: $r_1 = 2, \quad r_2 = 5, \quad r_3 = 5, \quad r_4 = 0, \quad r_5 = 0$
(het punt C, $w = 190$)
IV: $r_1 = 7, \quad r_2 = 2^1/_2, \quad r_3 = 0, \quad r_4 = 2^1/_2, \quad r_5 = 0$
(het punt B, $w = 215$).

Aldus is bij de vierde stap het maximum bereikt. De Simplex-techniek loopt van het ene hoekpunt naar een aangrenzend hoekpunt met een hogere waarde van de doelstellingsfunctie. Dit wordt geïllustreerd in de volgende figuur:



7. WAT IS DE PRIJS VAN EEN NEVENVOORWAARDE?

In het zojuist beschouwde voorbeeld van de radioproducent was de optimale oplossing $r_1 = 7$, $r_2 = 2^1/_2$ ($r_3 = 0$, $r_4 = 2^1/_2$, $r_5 = 0$). Bij een dergelijk productieprogramma wordt alle capaciteit op band 1 (voor eenvoudige radio's) gebruikt, de arbeid wordt volledig benut, doch de capaciteit van de tweede band (voor luxe radio's) wordt slechts voor 50% benut. *Als* de producent *nog* meer capaciteit op band 2 zou hebben gehad, zou de maximum te behalen winst niet zijn beïnvloed; er lagen

toch al $2^1/_2$ capaciteitseenheden braak. Eén extra eenheid capaciteit op band 2 verhoogt de maximum winst niet, en is dus voor de producent van geen waarde. Anders ligt de zaak voor de capaciteit op band 1. Indien de capaciteit van band 1 i.p.v. 7 radio's per dag 8 radio's per dag zou bedragen, dan zou de oplossing wel degelijk veranderen. Met behulp van een figuur als op blz. 24, waar nu echter de lijn AB één eenheid naar rechts is verschoven, kan men direct nagaan dat de nieuwe optimale oplossing aanleiding geeft tot het produceren van 8 eenvoudige radio's en 2 luxe radio's. (Men kan geen $2^1/_2$ luxe radio's blijven produceren i.v.m. de arbeidsrestrictie.) De totale winst wordt nu $8 \times 20 + 2 \times 30 = 220$. Dit is 5 gulden meer dan de 215 gulden in de oude oplossing, zodat wij concluderen dat *een eenheid capaciteit op band* 1 *een waarde van* 5 *gulden per dag heeft*. Ook arbeid is voor de producent iets waard. Een extra mandag immers maakt het mogelijk 7 eenvoudige en 3 luxe radio's te produceren. (Het kan wederom met behulp van een grafiek worden geverifieerd dat dit inderdaad de optimale oplossing is als er 13 mandagen arbeid beschikbaar zijn.) Dit productieprogramma geeft een winst van $7 \times 20 + 3 \times 30 = 230$; een extra winst van 15 gulden dus. Wij concluderen dat één mandag voor de producent een waarde heeft van 15 gulden. Als deze producent een werknemer voor minder dan 15 gulden per dag in dienst kan nemen, doet hij er verstandig aan dit te doen. (Het zal duidelijk zijn dat de waarde van een schaarse productiefactor voor een bedrijf niet altijd hoeft samen te vallen met de prijs ervan op de markt.)

Een analoge redenering kunnen wij houden voor het dieetprobleem. Wij moesten een zo goedkoop mogelijk dieet samenstellen dat aan een aantal minimum-vereisten betreffende calorieën e.d. voldeed. Men kan zich nu afvragen *hoeveel goedkoper* het goedkoopste dieet zou worden als we wat water in de wijn zouden doen, dus bijv. indien we met één calorie *minder* genoegen zouden nemen. Het is denkbaar dat het goedkoopste menu meer calorieën bevatte dan strikt noodzakelijk; dit gebeurt indien de *overige* voorwaarden betreffende vitaminen, eiwitten, enz. noopten tot het kopen van een menu dat het vereiste aantal calorieën overtreft. In zulk een geval zou men hetzelfde menu kopen zelfs als minder calorieën vereist zouden zijn; de calorie-voorwaarde verhoogt de kosten van het menu niet. Korter: *de restrictie kost niets*. Het is evenzeer mogelijk,

dat het goedkoopst mogelijke menu wél goedkoper wordt indien men één calorie minder nodig heeft. Wordt het menu dan b.v. één cent goedkoper, dan vormt dit een maatstaf voor de kosten van de restrictie.

De Simplex-techniek heeft nu de eigenschap, dat het tegelijk met de oplossing óók de prijs van de restricties geeft. Deze prijs wordt bepaald door het antwoord op de vraag: 'Indien wij één eenheid meer van het een of ander beschikbaar hebben, met hoeveel stijgt dan de waarde van de doelstellingsfunctie?' Of: 'Indien wij één eenheid minder van het een of ander nodig hebben, met hoeveel daalt dan de waarde van de doelstellingsfunctie?' (Een dergelijke prijs is steeds nul wanneer in de eindoplossing de spelingsvariabele van de desbetreffende restrictie een positieve waarde – niet nul – heeft; in dat geval is immers aan die restrictie ruimschoots voldaan en 'kost hij niets'.) Nadrukkelijk wijzen wij er op, dat het hier gaat om kleine toevoegingen of reducties, hier gemakshalve weergegeven met 'een eenheid meer of minder'. Anders kunnen er problemen rijzen, zoals men eenvoudig kan inzien. Stel, dat de radioproducent bijvoorbeeld 10 extra werknemers in dienst neemt voor in totaal 100 gulden. Hij denkt dat dit voordelig is, omdat hij redeneert dat 1 werknemer meer het mogelijk maakt $7 \times 20 + 3 \times 30 = 230$ gulden i.p.v. $7 \times 20 + 2^1/_2 \times 30 = 215$ gulden winst te maken, en daarom 15 gulden 'waard' is. Maar als het waar is dat één werknemer 15 gulden meer kan opbrengen, volgt daaruit dan noodzakelijkerwijs dat 10 werknemers extra 150 gulden meer zullen opbrengen? Neen, dit is niet zonder meer juist, *omdat het denkbaar is* (en bij grote veranderingen bepaald plausibel) dat nevenvoorwaarden waaraan aanvankelijk ruimschoots was voldaan bij grote veranderingen geschonden dreigen te worden. Zo ook hier. Zolang de nieuw aangenomen werknemer aan de lopende band van luxe radio's te werk kan worden gesteld, *zólang* blijft de extra opbrengst per werknemer met 15 gulden stijgen. Maar dit kan slechts tot op beperkte hoogte. Als er 5 werknemers meer zijn aangenomen is de 'luxe band' ook tot op capaciteit bezet! Beide banden zijn nu tot volle capaciteit bezet, en nóg meer werknemers verhogen de opbrengst niet. Het is in verband met dit soort problemen (dat nieuwe restricties opeens gaan 'klemmen') dat men slechts een beperkte hoeveelheid meer of minder moet beschouwen.

Naar wij zagen wordt een lineair programmeringsprobleem gekarakteriseerd door een drietal vereisten:

(i) Er moet een functie zijn die lineair is in een aantal beslissingsvariabelen; deze functie moet worden gemaximeerd of geminimeerd;

(ii) De beslissingsvariabelen mogen geen negatieve waarden aannemen;

(iii) Zij moeten voldoen aan een aantal lineaire nevenvoorwaarden, gewoonlijk in de vorm van ongelijkheden. (Soms is een aantal der nevenvoorwaarden al direct in de vorm van een gelijkheid gegeven. Wij kunnen ons dan de moeite van het invoeren van een spelingsvariabele besparen; het probleem wordt echter niet veranderd.)

Deze punten bepalen de structuur van het probleem. Hoe eenvoudig deze structuur ook moge zijn, zij laat ruimte voor een grote reeks van toepassingen op problemen van de meest verschillende aard. De moeilijkste opgave is veelal het *onderkennen* van een gegeven probleem als een lineair programmeringsvraagstuk. Het oplossen is daarna een, mogelijk bewerkelijke en tijdrovende, routinekwestie. Wij zullen twee voorbeelden geven, waarbij wij de nadruk leggen op het onderkennen en formuleren van het probleem.

Ons eerste voorbeeld heeft betrekking op het wedden op Duindigt. Het komt erop neer in een paardenrace dát paard aan te wijzen dat als eerste zal aankomen. Men kan dan wedden op dat paard, met een zeker bedrag als inzet. Wint dat paard inderdaad, dan wordt een zeker veelvoud van de inzet terugbetaald. Wint het paard de race niet, dan is de inzet verloren. Stel nu, dat er in een bepaalde race 4 paarden lopen: Wit, Bruin, Zwart en Schimmel. De uitbetaling geschiedt volgens de volgende koersen: Indien 1 gulden wordt ingezet op paard

Wit, en Wit wint, dan wordt 3 gulden uitbetaald;
Bruin, en Bruin wint, dan wordt 4 gulden uitbetaald;
Zwart, en Zwart wint, dan wordt 5 gulden uitbetaald;
Schimmel, en Schimmel wint, dan wordt 6 gulden uitbetaald.

In dit laatste geval (Schimmel wint) is de *winst* dus 6 gulden minus de gulden die werd ingezet, dus 5 gulden per saldo. Als het paard waarop men 1 gulden heeft gewed niet wint, dan is de inzet van 1 gulden verloren.

Gesteld nu, dat U 57 gulden hebt, hoeveel moet U dan op ieder paard wedden? Dat hangt, zo zult U denken, van Uw verwachtingen af. Als U er zeker van bent dat Bruin zal winnen, moet U natuurlijk 57 gulden op Bruin wedden. Dat is juist. *Doch deze redenering vóóronderstelt dat U zeker van Uw zaak bent.* En dat is niet vanzelfsprekend. In dit geval nemen wij aan dat U een volkomen vreemdeling bent op Duindigt, en een leek op paardengebied. Kortom: U hebt geen idee wie zal winnen. Wat dan? Misschien denkt U, als onverbeterlijk optimist, dat U, als U er toch niets van af weet, dan maar op Schimmel moet wedden: U hebt dan tenminste een kans om $57 \times 6 = 342$ gulden uitbetaald te krijgen voor 57 gulden inzet. Bij geen enkel ander paard hebt U een mogelijkheid van zo'n buitenkansje. Maar juist het feit dat U zoveel kunt verdienen als Schimmel wint, moet U argwanend maken; blijkbaar hebben de 'experts' niet veel vertrouwen in de capaciteiten van Schimmel. Volgens deze redenering is het wellicht beter op *Wit* te wedden; de 'experts' vinden een winst van Wit tenslotte het meest plausibel. Zij willen immers niet veel geld uitbetalen als Wit inderdaad wint. Maar ook als U op Wit wedt, hebt U de mogelijkheid 57 gulden te verliezen, terwijl U als U succes hebt 'slechts' $57 \times 3 = 171$ gulden terugkrijgt, dus $171 - 57 = 114$ gulden wint. En denkt U Uw gevoel van frustratie eens in wanneer U op Wit wedt en Schimmel wint tegen de verwachting in tòch? Had U $342 - 57 = 285$ gulden kunnen winnen en U hebt het niet gedaan!

Maar let nu eens op de volgende redenering. *Omdat U er geen idee van hebt wie zal winnen doet U er verstandig aan U te dekken tegen de ongunstigste mogelijkheid.* Pessimistisch? Conservatief? Een beetje wel. Maar als U Uw huis verzekert tegen brand, dekt U zich immers ook tegen de meest ongunstige samenloop van omstandigheden? Deze filosofie zal onze handelwijze bepalen; het zal leiden tot een heel fraai lineair programmeringsprobleem.

De *beslissingsvariabelen* in dit probleem zijn de bedragen die worden ingezet op Wit, Bruin, Zwart en Schimmel. Voor deze bedragen zullen wij kortheidshalve schrijven x_W, x_B, x_Z en x_S. Zij mogen niet negatief zijn, en moeten (in dit voorbeeld) voldoen aan de gelijkheid

$$x_W + x_B + x_Z + x_S = 57,$$

want de totale inzet is 57 gulden. De *doelstellingsfunctie* moet

erop gericht zijn er in het ongunstigste geval zo goed mogelijk af te komen; dus om de *minimum* winst ('het ongunstigste geval') te *maximeren* ('er zo goed mogelijk af komen'). Uiteraard is het denkbaar, dat de hoogste minimum winst negatief is, dus op een verlies neerkomt.

Hoe hoog is de winst als besloten wordt x_W, x_B, x_Z, x_S op de vier paarden in te zetten? Dat hangt ervan af welk paard wint! Als Wit wint brengen alleen de x_W guldens op Wit ingezet hun geld op, nl. $3x_W$ gulden omdat het tarief voor Wit 3 bedraagt. Na aftrek van de totale inzet ad 57 gulden blijft dus een winst van $3x_W - 57$ over (die natuurlijk negatief is, dus een verlies, als x_W kleiner is dan 19). Wanneer we dit herhalen voor de drie andere mogelijkheden komen we tot het volgende schema:

Als Wit wint, bedraagt de winst $3x_W - 57$;
als Bruin wint, bedraagt de winst $4x_B - 57$;
als Zwart wint, bedraagt de winst $5x_Z - 57$;
als Schimmel wint, bedraagt de winst $6x_S - 57$.

Het gaat erom, het minimum (M) van deze vier bedragen zo groot mogelijk te doen zijn. Hoe drukken we M in onze beslissingsvariabelen uit? Wel, door op te merken dat de vier winsten elk tenminste gelijk zijn aan M – dat volgt uit de definitie van het minimum! Dus:

$$3x_W - 57 \geqq M$$
$$4x_B - 57 \geqq M$$
$$5x_Z - 57 \geqq M$$
$$6x_S - 57 \geqq M.$$

Of ook, met behulp van niet-negatieve spelingsvariabelen (y_W, y_B, y_Z, y_S):

$$3x_W - 57 - y_W = M$$
$$4x_B - 57 - y_B = M$$
$$5x_Z - 57 - y_Z = M$$
$$6x_S - 57 - y_S = M.$$

Tel nu deze vier vergelijkingen op:

$$3x_W + 4x_B + 5x_Z + 6x_S - 228 - y_W - y_B - y_Z - y_S = 4M.$$

Wij kunnen dit resultaat nog iets vereenvoudigen door rekening te houden met het feit, dat de som van de x-en gelijk is aan 57. Na vermenigvuldiging met 4:

$$4x_W + 4x_B + 4x_Z + 4x_S = 228.$$

Vullen we dit in de zojuist verkregen vergelijking in, dan vinden we:

$$4M = -x_W + x_Z + 2x_S - y_W - y_B - y_Z - y_S,$$

zodat x_B is uitgevallen. Het resultaat voor M is dus:

$$M = {}^1/_4(-x_W + x_Z + 2x_S - y_W - y_B - y_Z - y_S),$$

en dat is de doelstellingsfunctie die gemaximeerd moet worden.

Dit is een lineair programmeringsvraagstuk! We hebben immers een doelstellingsfunctie die lineair is in 7 variabelen die alle niet-negatief dienen te zijn; en ook de nevenvoorwaarden zijn lineair.[1] Hoewel het ons meer om de formulering van het vraagstuk gaat dan om de oplossing is het wel aardig te vermelden, dat de optimale oplossing de volgende is:

$x_W = 20$ (dus is de winst $3x_W - 57 = 60 - 57$ als Wit wint)
$x_B = 15$ (dus is de winst $4x_B - 57 = 60 - 57$ als Bruin wint)
$x_Z = 12$ (dus is de winst $5x_Z - 57 = 60 - 57$ als Zwart wint)
$x_S = 10$ (dus is de winst $6x_S - 57 = 60 - 57$ als Schimmel wint).

Als men 57 gulden op *deze* wijze (20/15/12/10) uitsplitst, is men er *zeker* van, *wát* er ook gebeurt, *welk* paard ook wint, 60 gulden te ontvangen. En dus $60 - 57 = 3$ gulden te winnen; dus is de maximale M gelijk aan 3. Iedere onzekerheid is verdwenen. U kunt dit vergelijken met een brandverzekering voor Uw huis van 30.000 gulden. U ruilt dan het risico van een verlies van 30.000 gulden in tegen de zekerheid van het verlies van de verzekeringspremie. In het hoofdstuk over speltheorie zullen wij hierop terugkomen.

9. EEN FABRIKANT

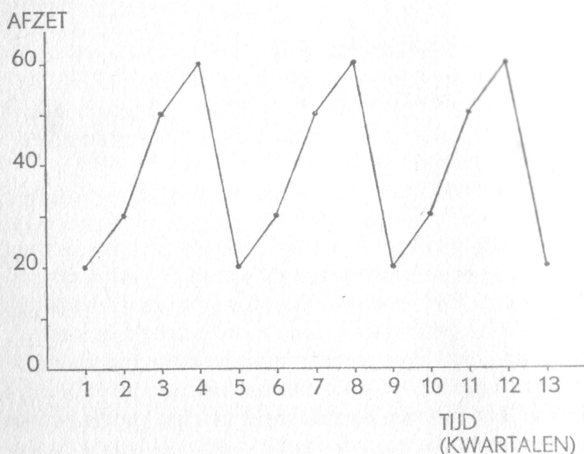
Ons tweede voorbeeld betreft een fabrikant, wiens doel het is zo goedkoop mogelijk te produceren. Om de zaak niet te ge-

1. Er zijn er vier. De eerste zegt, dat de som van alle x-en gelijk is aan 57. De drie overige impliceren, dat
$3x_W - 57 - y_W$, $4x_B - 57 - y_B$, $5x_Z - 57 - y_Z$, $6x_S - 57 - y_S$
alle aan elkaar gelijk zijn. (Zij zijn immers alle gelijk aan M!)

compliceerd te maken zullen we ons tot een enkel product bepalen en ervan uitgaan dat de afzet per kwartaal van tevoren vast ligt:

20 in het eerste kwartaal (in zekere eenheden)
30 in het tweede kwartaal
50 in het derde kwartaal
60 in het vierde kwartaal.

En daarna weer van voren af aan, dus 20 in het vijfde, 30 in het zesde kwartaal, enz. Er is dus een jaarlijks seizoenpatroon over de kwartalen, dat stabiel in het verloop van de tijd is. In een figuur:



De fabrikant moet dus ieder jaar $20 + 30 + 50 + 60 = 160$ eenheden produceren. De vraag is *hoe hij deze productie over de verschillende kwartalen moet verdelen, als hij zijn productiekosten wil minimeren.* De doelstelling is dus: het minimeren van de totale productiekosten. De beslissingsvariabelen zijn: de productieomvang in ieder kwartaal. Deze beslissingsvariabelen zullen wij aangegeven met $x(t)$; dus $x(1)$ is de productieomvang in het eerste kwartaal; $x(2)$ is de productieomvang in het tweede kwartaal; . . . $x(17)$ is de productieomvang in het zeventiende kwartaal. Maar het probleem is 'periodiek'; het is ieder jaar

opnieuw hetzelfde. Daarom is $x(1) = x(5)$; $x(2) = x(6)$; $x(3) = x(7)$; enz. Er zijn dus eigenlijk slechts vier beslissingsvariabelen: $x(1)$, $x(2)$, $x(3)$ en $x(4)$. Wij merken al direct op dat uit de aard van het probleem volgt dat die variabelen geen negatieve waarden kunnen aannemen.

Nu de doelstellingsfunctie: de totale productiekosten per jaar. Het is in dit verband van belang, dat we met twee soorten van productiekosten te doen hebben. De ene categorie omvat de vaste lasten (bijv. van het fabrieksgebouw, dat er al staat) en ook de kosten die rechtstreeks verbonden zijn aan het feit dat er nu eenmaal 160 eenheden per jaar geproduceerd moeten worden. Aan deze categorie kan de fabrikant niets doen met behulp van de hier beschouwde beslissingsvariabelen; die kunnen we dus gevoegelijk buiten beschouwing laten. De andere categorie wordt wèl beïnvloed door deze beslissingsvariabelen en valt in twee soorten uiteen:

(i) *Voorraadkosten.* Voor alle eenheden die aan het einde van een kwartaal in voorraad zijn (dus wél geproduceerd maar niet verkocht zijn), heeft de fabrikant 8 gulden kosten. De beginvoorraad, dus de voorraad aan het begin van kwartaal 1, is nul.[1]

(ii) *Instelkosten.* Aan het begin van ieder kwartaal t moet de apparatuur worden *ingesteld* op de productieomvang van dat kwartaal, $x(t)$. De fabrikant heeft voor iedere eenheid verschil in de productieomvang van twee *opeenvolgende* kwartalen 5 gulden kosten. Dus als $x(3) = 50$, en $x(4) = 40$, dan heeft de fabrikant $(50 - 40) \times 5 = 50$ gulden kosten. Die kosten zijn precies dezelfde in het omgekeerde geval, dus wanneer de productie stijgt van $x(3) = 40$ tot $x(4) = 50$.

Laten we ons nu voorstellen, dat de fabrikant twee jonge employé's heeft aangetrokken, een econoom en een ingenieur, en dat hij beiden vraagt op welke niveau de successieve kwartaalproducties gesteld dienen te worden. Tijd om te rekenen is er niet, het antwoord moet meteen komen. Dan openbaren zich de verschillen in opleiding. De econoom weet van de vele jaren die hij op de collegebanken doorbracht dat voorraad houden geld kost. Hij adviseert per kwartaal precies evenveel te produceren als met de vraag in dat kwartaal overeenstemt. De inge-

1. Het is een wat bizarre consequentie van de veronderstelling dat de afzet volledig bepaald is, dat er geen "ijzeren" of "buffer"-voorraad nodig is.

nieur heeft even lang op de collegebanken doorgebracht, maar hij is op andere wijze geconditioneerd: hij weet dat het voor de productiechef bijzonder prettig is constante hoeveelheden te produceren en adviseert dienovereenkomstig.

Wat zijn de kostenimplicaties van deze strijdige adviezen? Het eerste advies, dat van de econoom, vermijdt de voorraadkosten, maar het leidt tot hoge instelkosten. Deze kan men als volgt bepalen. Van kwartaal 1 naar kwartaal 2 stijgt de afzet en dus ook (onder de hier gemaakte veronderstelling) de productie met 10 eenheden, nl. van 20 tot 30. De instelkosten bedragen dus 10 × 5 = 50 gulden. Van kwartaal 2 naar 3 is er een stijging van 20 eenheden (30 naar 50), dus 100 gulden instelkosten. Van kwartaal 3 naar 4 is er een stijging van 10 eenheden (50 naar 60), dus 50 gulden instelkosten. Van kwartaal 4 naar 5 is er een daling van 40 eenheden (van 60 terug naar 20), dus 200 gulden instelkosten. Daarmee is het jaar rond. In totaal dus:

50 + 100 + 50 + 200 = 400 gulden

instelkosten per jaar. En aangezien de voorraadkosten vermeden zijn, is dit dus tevens het bedrag van de totale productiekosten (voorzover we die hier beschouwen, dus instel- en voorraadkosten bij elkaar geteld).

Het advies van de ingenieur komt erop neer, dat de instelkosten vermeden worden door per kwartaal een constante hoeveelheid te produceren. Aangezien per jaar 160 eenheden nodig zijn komt dit neer op 40 eenheden per kwartaal. We hebben dan geen instelkosten maar wel (hoge) voorraadkosten. Om die te bepalen moeten we eerst de voorraad aan het einde van elk kwartaal berekenen. Nu is de beginvoorraad (aan het begin van het eerste kwartaal) nul. Gedurende het eerste kwartaal worden 40 eenheden geproduceerd en 20 verkocht; het restant, dus 20 eenheden, is dan de voorraad aan het einde van het eerste kwartaal. Dit is tevens ook de voorraad aan het begin van het tweede kwartaal. Gedurende dat kwartaal worden ook weer 40 eenheden geproduceerd, maar nu worden er 30 verkocht; het restant is dus slechts 10. Dit komt bij de voorraad die er aan het begin van het tweede kwartaal al was, dus 20 + 10 = 30 is de voorraad aan het einde van het tweede kwartaal. Enzovoorts. In het derde en vierde kwartaal wordt minder geproduceerd (40) dan verkocht (50 resp. 60), zodat de voorraad dan gaat dalen; op het eind van het vierde kwartaal is de voor-

raad weer nul, net als bij het begin, en daarna gaat het weer van voren af aan. Een volledig overzicht voor de eerste vier kwartalen vindt men in het volgende schema:

Kwartaal	Afzet	Prod.	Voorraad einde kwartaal (beginvoorr. + prod. — afzet)
1	20	40	$0 + 40 - 20 = 20$
2	30	40	$20 + 40 - 30 = 30$
3	50	40	$30 + 40 - 50 = 20$
4	60	40	$20 + 40 - 60 = 0$

De eindvoorraad van de vier successieve kwartalen is dus 20, 30, 20 resp. 0. De kosten per eenheid voorraad bedragen 8 gulden, dus hebben we hier de volgende totale voorraadkosten per jaar:

$$8 \times 20 + 8 \times 30 + 8 \times 20 + 8 \times 0 = 560 \text{ gulden.}$$

Dit is tevens het bedrag van de totale kosten, omdat we de instelkosten immers vermeden hebben.

Dit zijn de uitersten; en we kunnen het redelijke vermoeden hebben, dat het optimum ergens tussen de twee adviezen in zal liggen. Dan hebben we niet uitsluitend instelkosten en ook niet uitsluitend voorraadkosten; we hebben beide kostensoorten tegelijkertijd, maar hun *som* (het totaal) zal dan geringer moeten zijn dan de 400 gulden van het ene uiterste, laat staan de 560 gulden van het andere. Om dit te bereiken moeten wij allereerst de instel- en voorraadkosten in de beslissingsvariabelen uitdrukken. Wij beginnen met de voorraadkosten en gaan daarbij te werk volgens het hierboven gegeven schema. Aan het begin van het eerste kwartaal is de voorraad nul. Gedurende dat kwartaal wordt $x(1)$ geproduceerd [in het schema hadden we het bijzondere geval $x(1) = 40$] en wordt 20 verkocht. Dus:

Voorraad aan het einde van het eerste kwartaal
$= 0 + x(1) - 20 = x(1) - 20.$

Dat is tevens de voorraad aan het begin van het tweede kwartaal. Om de voorraad aan het einde van dat kwartaal te vinden moeten we de productie van dat kwartaal, $x(2)$, bijtellen en de afzet (30) aftrekken. Dus:

Voorraad aan het einde van het tweede kwartaal
$= x(1) - 20 + x(2) - 30 = x(1) + x(2) - 50,$

en op dezelfde manier is ook dit weer de voorraad aan het begin van het derde kwartaal. Daar tellen we de productie van dat kwartaal bij, $x(3)$, en trekken de afzet (50) af teneinde de voorraad aan het einde van het kwartaal te vinden:

Voorraad aan het einde van het derde kwartaal
$= x(1) + x(2) - 50 + x(3) - 50 = x(1) + x(2) + x(3) - 100.$

Volkomen analoog voor het vierde en laatste kwartaal:

Voorraad aan het einde van het vierde kwartaal
$= x(1) + x(2) + x(3) - 100 + x(4) - 60$
$= x(1) + x(2) + x(3) + x(4) - 160.$

Maar die laatste voorraad is noodzakelijk nul, want aan het einde van het vierde kwartaal zijn we weer aan het begin van het eerste kwartaal van het volgende jaar! Dat ziet men ook onmiddellijk aan de hand van de zojuist afgeleide formule, want de som van de vier x-en, $x(1) + \ldots + x(4)$, is noodzakelijkerwijs gelijk aan 160 als de fabrikant precies aan de jaarlijkse vraag wil voldoen.

Nu kost een eenheid voorraad per kwartaal 8 gulden. We moeten dus de drie voorraden die niet nul zijn bij elkaar optellen en met 8 vermenigvuldigen om tot de totale voorraadkosten per jaar te komen:

Totale voorraadkosten per jaar in guldens
$= 8[x(1) - 20 + x(1) + x(2) - 50 + x(1) + x(2) + x(3) - 100]$
$= 24x(1) + 16x(2) + 8x(3) - 1360.$

Aldus zijn de voorraadkosten in de beslissingsvariabelen uitgedrukt. Vervolgens de instelkosten. Deze worden bepaald door de successieve verschillen van de productieomvang in de diverse kwartalen, dus door

$x(2) - x(1), \quad x(3) - x(2), \quad x(4) - x(3), \quad x(5) - x(4).$

Het is van belang eerst het teken (plus, nul of min) van deze

successieve verschillen te bestuderen. Kunnen wij bijv. iets zeggen over het teken van $x(2) - x(1)$, dus over de vraag of de productie in het tweede kwartaal groter of kleiner dan die van het eerste kwartaal dient te zijn? Neem het geval $x(1) = 40$, $x(2) = 20$, zodat er dan in het eerste kwartaal meer wordt geproduceerd dan in het eerste; en laten we dit geval vergelijken met $x(1) = x(2) = 30$, waarin dus in beide kwartalen *tezamen* evenveel wordt geproduceerd als in het eerste geval, maar waarin de productie van het tweede kwartaal niet lager is dan die van het eerste kwartaal. Kennelijk is het tweede geval voordeliger dan het eerste wat de instelkosten betreft, want die zijn dan nul. Maar het is òòk voordeliger wat de voorraadkosten betreft. Om dit in te zien moeten we bedenken, dat de voorraad aan het einde van het tweede kwartaal in beide gevallen dezelfde is. We vonden immers dat die gelijk is aan $x(1) + x(2) - 50$, dus hier 10 in beide gevallen. Het gaat dus om de voorraad aan het einde van het eerste kwartaal, waarvan wij vonden dat die gelijk is aan $x(1) - 20$, d.w.z. $x(1)$ minus de afzet van het eerste kwartaal. Die voorraad is kennelijk groter in het eerste geval (20) dan in het tweede (10) en brengt dienovereenkomstig ook grotere kosten met zich mee. De lezer zal aanvoelen, dat wat de instelkosten betreft het voordelig is de productieomvang niet te veranderen (althans zo min mogelijk te veranderen); en dat het t.a.v. de voorraadkosten aanbeveling verdient de productie zoveel mogelijk gelijk op met de afzet te doen bewegen. Niet tegen de draad in! En dat is wat het geval $x(1) = 40$, $x(2) = 20$ precies doet, want de afzet klimt van 20 naar 30 in dezelfde periode.

Op grond van deze overwegingen (die eigenlijk nog wel iets exacter zouden moeten zijn) komen we tot de conclusie dat in de oplossing $x(2)$ niet lager zal zijn dan $x(1)$. Ligt $x(2)$ hoger, dan moeten instelkosten worden gemaakt ten bedrage van $5[x(2) - x(1)]$, want het tarief is 5 gulden per eenheid verschil. Op overeenkomstige wijze zijn de instelkosten van de twee volgende kwartalen:

$5[x(3) - x(2)]$ resp. $5[x(4) - x(3)]$,

want de afzet klimt in die kwartalen en de productie zal zich daaraan aanpassen: $x(3) \geqq x(2)$, $x(4) \geqq x(3)$. Daarna, in het vijfde kwartaal, daalt de afzet scherp: van 60 naar 20. Dus zal ook de productie dalen en bedragen de instelkosten $5[x(4) - x(5)]$, niet $5[x(5) - x(4)]$. Wanneer we nog in herinnering

brengen dat $x(5)$ hetzelfde is als $x(1)$, dan vinden we na optelling:

Totale instelkosten per jaar in guldens
$= 5[x(2) - x(1) + x(3) - x(2) + x(4) - x(3) + x(4) - x(1)]$
$= -10x(1) + 10x(4)$.

De totale productiekosten (voorraad plus instel) bedragen dus:
$K = 14x(1) + 16x(2) + 8x(3) + 10x(4) - 1360.$
Dit is een lineaire doelstellingsfunctie die moet worden geminimeerd. (Het getal —1360 is een constante die in het eigenlijke minimeringsproces geen rol speelt, omdat hij toch niet beïnvloed kan worden.)

Om te laten zien, dat we inderdaad met een lineair programmeringsvraagstuk in vier beslissingsvariabelen te doen hebben, moeten we nog de restricties beschouwen. Allereerst zijn er de niet-negativiteitsrestricties voor elk van deze variabelen. Vervolgens de restrictie dat hun som gelijk is aan 160; dat is dus een lineaire vergelijking. Tenslotte, wanneer we ervan uitgaan dat de voorraden niet negatief mogen zijn, hebben we ook nog de volgende restricties:

$x(1) \geqq 20$
$x(1) + x(2) \geqq 50$
$x(1) + x(2) + x(3) \geqq 100,$

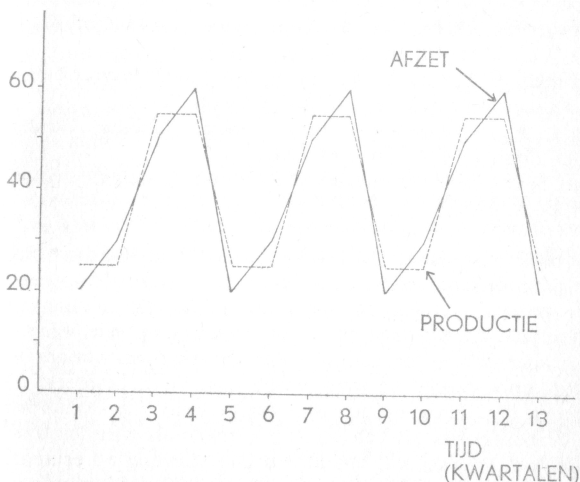
zoals men gemakkelijk verifieert aan de hand van de voorraadformules van de voorgaande bladzijden. Dit zijn lineaire ongelijkheden, zodat we inderdaad met een lineair programmeringsvraagstuk te doen hebben.

Volledigheidshalve vermelden we de optimale oplossing: produceer 25 eenheden in de eerste twee kwartalen, 55 eenheden in de laatste twee. Dit betekent, dat er per jaar slechts twee keer instelkosten worden gemaakt: aan het eind van het tweede kwartaal en aan het eind van het vierde kwartaal. Kosten per keer: 5 × 30 = 150 gulden. De totale instelkosten bedragen dus 300 gulden per jaar. De voorraadstaat:

Kwartaal	Afzet	Prod.	Voorraad einde kwartaal (beginvoorr. + prod. — afzet)
1	20	25	0 + 25 — 20 = 5
2	30	25	5 + 25 — 30 = 0

3	50	55	0 + 55 — 50 = 5
4	60	55	5 + 55 — 60 = 0.

De voorraadkosten bedragen dus 8(5 + 0 + 5 + 0) = 80 gulden per jaar. Totale productiekosten 300 + 80 = 380 gulden per jaar. Dit ligt 20 gulden onder het kostenbedrag verbonden aan het advies van de econoom, dat neerkwam op onmiddellijke aanpassing van de productie aan de afzet. In de volgende figuur vindt men het verloop van productie en afzet beide:



LITERATUUR

De Simplex-techniek is afkomstig van G.B. Dantzig [1]. De techniek was al eerder bij ingewijden bekend door verschillende rapporten van dezelfde schrijver (soms samen met M. K. Wood). Trouwens, de Rus L. Kantorovich had al in 1939 een verwante oplossingstechniek gegeven. Hij was ook de eerste, die een 'echt' lineair programmeringsvraagstuk formuleerde, hetgeen overigens buiten Rusland pas omstreeks 1955 werd onderkend. Zijn oorspronkelijke publicatie is vertaald [2]. Het dieetprobleem, dat herhaaldelijk in dit hoofdstuk werd vermeld, is ontleend aan Stigler [3]. Het probleem inzake de omvang van de productie, dat aan het einde van dit hoofdstuk werd behandeld, is voor het eerst gesteld in [4]; sindsdien is er nog herhaaldelijk aandacht aan besteed. Het Duindigt-probleem is voorzover

ons bekend het eerst gesteld door Vajda [5]. Vajda's boek kan dienen als nadere inleiding tot lineair programmeren. Een iets eenvoudiger boek, dat zeer helder geschreven is, is dat van Gass [6].

[1] Dantzig, G. B., 'Maximization of a Linear Function of Variables Subject to Linear Inequalities'. Hoofdstuk XXI van *Activity Analysis of Production and Allocation*, edited by T. C. Koopmans. John Wiley and Sons, Inc., New York. 1951.

[2] Kantorovich, L., 'Mathematical Methods of Organizing and Planning Production'. *Management Science*, Vol. 6 (1960), pp. 366-422.

[3] Stigler, G. J., 'The Cost of Subsistence'. *Journal of Farm Economics*, Vol. 27 (1945), pp. 303-314.

[4] Hoffman, A. J., and W. W. Jacobs, 'Smooth Patterns of Production'. *Management Science*, Vol. 1 (1954), pp. 86-91.

[5] Vajda, S., *Mathematical Programming*. Addison Wesley Publishing Company, Inc., London. 1961.

[6] Gass, S. I., *Linear Programming*. McGraw-Hill Book Company, Inc., New York. 1958.

Zowel het boek van Gass als dat van Vajda vermelden veel meer dan lineair programmeren, de Simplex-techniek, en een aantal voorbeelden. Gass vermeldt uitvoerig de moeilijkheden die kunnen ontstaan door *degeneratie*. Degeneratie treedt op wanneer op een gegeven moment *minder* dan m (het aantal nevenvoorwaarden) variabelen positief zijn, i.p.v. precies m, zoals gewoonlijk. Vajda bespreekt de moeilijkheden die zich voordoen wanneer de oplossing slechts zin heeft als de beslissingsvariabelen *gehele getallen* als waarden aannemen. Dit kan voorkomen, omdat een half vliegtuig nu eenmaal niet kan vliegen. Ook gaat Vajda in op *kwadratisch* programmeren. Hier spreekt men van wanneer de doelstellingsfunctie kwadratisch is in de beslissingsvariabelen; overigens pleegt men ook dan met restricties in de vorm van lineaire ongelijkheden te werken. Een kwadratische functie in de beslissingsvariabelen x_1 en x_2 zou bijv. kunnen luiden: $-4x_1^2 - x_2^2 + 2x_1x_2 + 6x_1 + 4x_2$.

Beide boeken geven naast de Simplex-techniek een aantal andere oplossingstechnieken. Sommige hiervan zijn bijzonder efficiënt voor speciale problemen; bijvoorbeeld voor transportproblemen, die wij in Hoofdstuk 2 uitvoeriger zullen bespreken. Beide boeken bevatten een uitgebreide literatuuropgave.

2. HET OPTIMALE EN HET KRITIEKE PAD

1. BRAND

Wij beginnen dit hoofdstuk met het schetsen van het decor van een belangrijke toepassing van lineair programmeren. Het betreft hier een transportprobleem. Hoewel dit soort problemen met behulp van de gebruikelijke Simplex-techniek wel is op te lossen, zijn er snellere, efficiëntere methoden. Deze zullen wij nu uitvoerig en in detail behandelen.

In de stad Rotterdam (het voorbeeld is fictief) zijn drie brandweergarages: één in Zuid, één in Oost en één in West. Deze hebben respectievelijk 2, 3 en 4 brandweerwagens, die bij alarm binnen een halve minuut kunnen uitrukken. Voor licht-alarm is 1 wagen voldoende; voor middel-alarm zijn 3 wagens voldoende; voor groot-alarm rukt men in eerste instantie met 5 wagens uit (welke dus niet alle uit dezelfde garage kunnen komen, omdat zelfs de grootste garage slechts 4 auto's bevat).

Verder is er een brandweerhoofdman. Deze zit achter een bureau op het stadhuis. Op het bureau staan drie telefoons. Een *groene* voor licht-alarm, een *witte* voor middel-alarm, en een *rode* voor groot-alarm. Het is de taak van de brandweerhoofdman bij het binnenkomen van meldingen de brandweerauto's uit de garages naar de brand te dirigeren. Hij zal daarbij proberen *die* garage te bellen welke het dichtst bij de brand is, zodat de auto's zo snel mogelijk aanwezig zullen zijn.

Op een gegeven moment – alsof de duivel ermee speelt – rinkelen de drie telefoons van de brandweerhoofdman tegelijkertijd. De groene meldt een licht-alarm in het centrum; de witte een middel-alarm in Zuid, en de rode een groot-alarm in Noord. De vraag voor de hoofdman is nu de volgende: *hoe* moeten de wagens uit de verschillende garages naar de respectieve brandhaarden worden gedirigeerd, zodanig dat het aantal kilometers door de 9 brandweerauto's gezamenlijk af te leggen minimaal is? Het betreft hier dus een minimeringsprobleem. Bovendien zijn er nevenvoorwaarden, die wij zullen formuleren met behulp van het volgende overzichtelijke schema:

	(1) Brand Centrum	(2) Brand Zuid	(3) Brand Noord
	1	3	5
(1) Garage Zuid 2	5	2	9
(2) Garage Oost 3	3	8	6
(3) Garage West 4	2	5	7

In dit schema staan in de linkerkolom de in de verschillende garages beschikbare auto's (2, 3, 4); in de bovenrij staat het aantal auto's vereist bij de verschillende branden (1, 3, 5). In de tabel zélf staan de afstanden; de afstand tussen garage Oost en brand Noord bedraagt 6 kilometer; de afstand tussen garage West en de brand in Zuid bedraagt 5 kilometer.

De *beslissingsvariabelen* in dit probleem zijn het *aantal* auto's dat van een bepaalde garage naar een bepaalde brand wordt gezonden. Om de notatie handzaam te houden noemen wij de garages Zuid, Oost en West respectievelijk 1, 2 en 3 (zoals in het schema al is aangegeven). Ook de branden Centrum, Zuid en Noord geven wij aan met 1, 2 resp. 3. Nu schrijven wij x_{11} voor het aantal auto's dat van Zuid naar Centrum wordt gezonden. De *eerste* onderindex geeft aan *vanwaar* (van welke garage) de auto wordt verzonden; de *tweede* onderindex geeft aan *waarheen* (naar welke brand) de auto wordt gezonden. Evenzo staat x_{23} voor het aantal auto's dat van 2 (garage Oost) naar 3 (brand Noord) wordt gezonden.

De doelstellingsfunctie laat zich nu m.b.v. de 9 variabelen en de gegevens uit het schema eenvoudig formuleren. Wij zien immers onmiddellijk, dat de x_{11} wagens die van 1 naar 1 rijden elk 5 kilometer afleggen, dus $5x_{11}$ in totaal; dat de x_{12} wagens die van 1 naar 2 gaan elk 2 kilometer rijden, dus $2x_{12}$ in totaal; enzovoorts, en al die afstanden moeten worden opgeteld. De doelstellingsfunctie is dan het totaal aantal autokilometers,

$$\begin{aligned} &5x_{11} + 2x_{12} + 9x_{13} \\ &+ 3x_{21} + 8x_{22} + 6x_{23} \\ &+ 2x_{31} + 5x_{32} + 7x_{33}, \end{aligned}$$

en dit moet worden geminimeerd. Ook de nevenvoorwaarden kunnen eenvoudig worden geformuleerd. Alle auto's zullen hun garage moeten verlaten, want er zijn 9 auto's nodig bij de verschillende branden, en er staan ook 9 auto's in de garages. Wij hebben nu als restricties:

(1) $x_{11} + x_{12} + x_{13} = 2$
[het aantal auto's gezonden uit 1 (garage Zuid) is 2]

(2) $x_{21} + x_{22} + x_{23} = 3$
[het aantal auto's gezonden uit 2 (garage Oost) is 3]

(3) $x_{31} + x_{32} + x_{33} = 4$
[het aantal auto's gezonden uit 3 (garage West) is 4].

Dit zijn de voorwaarden die het gevolg zijn van het aantal wagens in de garages aanwezig. Daarnaast hebben wij soortgelijke restricties in verband met het aantal bij de verschillende branden vereiste auto's:

(4) $x_{11} + x_{21} + x_{31} = 1$
[het aantal naar 1 (brand Centrum) te zenden auto's is 1]

(5) $x_{12} + x_{22} + x_{32} = 3$
[het aantal naar 2 (brand Zuid) te zenden auto's is 3]

(6) $x_{13} + x_{23} + x_{33} = 5$
[het aantal naar 3 (brand Noord) te zenden auto's is 5].

Tenslotte mogen de beslissingsvariabelen niet negatief zijn. Het is dus een duidelijk lineair programmeringsprobleem, waarbij echter de nevenvoorwaarden al *direct* in de vorm van gelijkheden (vergelijkingen) i.p.v. ongelijkheden zijn geformuleerd. Dit brengt geen enkele complicatie met zich mee. We hebben immers in het vorige hoofdstuk gezien, dat men ongelijkheden juist transformeert tot gelijkheden m.b.v. spelingsvariabelen; dat is hier niet nodig, spelingsvariabelen zijn dus overbodig.

Het lijkt alsof onze 9 beslissingsvariabelen aan 6 nevenvoorwaarden dienen te voldoen, maar dat is toch niet helemaal correct. Wij kunnen nl. één van de nevenvoorwaarden uit de 5 overige afleiden, zodat er dus in feite maar 5 nevenvoorwaarden zijn. Dit vloeit voort uit de omstandigheid dat er precies evenveel auto's in de garages staan (2 + 3 + 4) als er nodig zijn om de branden te blussen (1 + 3 + 5). Wanneer we voor alle drie garages zouden specificeren hoeveel van hun auto's gaan naar brand 1 en 2; en bovendien voor garages 1 en 2

hoeveel er gaan naar brand 3; dan ziet men gemakkelijk, dat het niet anders kan dan dat de resterende auto's van garage 3 naar brand 3 gaan. Dat is eenvoudig het sluitstuk. Wiskundig ziet men dit ook heel simpel: tel de eerste drie nevenvoorwaarden op en trek dan de vierde en de vijfde af – wat overblijft is de zesde nevenvoorwaarde! Anders gezegd, we kunnen ons tot de eerste 5 nevenvoorwaarden bepalen, aan de zesde is dan automatisch voldaan. (We kunnen ook bijv. de laatste 5 i.p.v. de eerste 5 nevenvoorwaarden nemen, maar we zullen het op deze keus houden om de aandacht te bepalen.)[1]

We hebben dus te doen met een lineair programmeringsprobleem in 9 variabelen en 5 nevenvoorwaarden. Sinds het vorige hoofdstuk weten we, dat we voor de oplossing dan dienen te zoeken naar 5 positieve variabelen; het restant is nul. Echter, als de brandweerhoofdman op het stadhuis dit probleem met de gebruikelijke (Simplex) techniek zou trachten op te lossen, zouden de brandende en belendende percelen zijn afgebrand vóór er ook maar één auto zijn garage zou hebben verlaten. Vandaar dat een snellere techniek is vereist. Deze techniek zal bovendien leiden tot het inzicht dat de oplossing altijd zal bestaan uit gehele getallen! Dit is een gelukkige bijkomstigheid. Immers, als het lineair programmeringsvraagstuk als deel van de oplossing bijv. zou geven dat er $1^1/_2$ auto van garage Zuid naar brand Zuid moet, en de resterende $^1/_2$ auto van garage Zuid naar brand Centrum, dan waren we weinig verder. Dit nu kan niet voorkomen.

Nog even dit. Wij construeerden een wat gezocht voorbeeld van een zgn. transportprobleem. Men moet echter onderkennen dat de problematiek veel algemener is. Bijvoorbeeld: hoeveel Shell tankschepen zullen van welke Shell olie-raffinaderijen hoeveel olie waarheen vervoeren? In het algemeen, overal waar verschillende productiecentra en distributiecentra zijn, die door één instantie worden beheerd, doet zich dit type van transportproblemen voor.

1. De zaak loopt iets anders, wanneer er meer auto's in de garages staan dan nodig zijn om de branden te blussen. Dit valt overigens vrij eenvoudig m.b.v. spelingsvariabelen op te lossen (die dan de betekenis hebben van het aantal auto's dat in een bepaalde garage gehouden wordt), maar we beperken ons hier tot het eenvoudigste geval.

Een van de grote voordelen van een transportprobleem is dat het bijzonder eenvoudig is een oplossing te vinden die (i) aan alle nevenvoorwaarden voldoet, en (ii) precies evenveel positieve variabelen heeft als er nevenvoorwaarden zijn. Vooropgesteld, dat deze nevenvoorwaarden op de juiste wijze worden geteld, dus één minder dan de som van het aantal garages (of productiecentra) en het aantal branden (of distributiecentra). Wij kunnen dit aan de hand van het brandweervoorbeeld illustreren. Daartoe maken wij weer een schema met als linkerkolom de beschikbare auto's, en als bovenrij het vereiste aantal auto's:

	1	3	5
2	.	.	.
3	.	.	.
4	.	.	.

De getallen die wij nu op de plaats van de stippen zullen invullen zijn niet, zoals in het schema op blz. 50, kilometerafstanden, maar aantallen auto's die van een garage (behorend bij de rij) naar een brand (behorend bij een kolom) zullen worden gereden. Wij gaan te werk volgens hetgeen in de Angelsaksische literatuur bekend staat als de 'Northwest corner rule'. D.w.z., wij beginnen linksboven, dus bij garage 1 en brand 1; stap voor stap gaan we naar rechts en omlaag tot we uiteindelijk rechtsonder (de laatste garage en de laatste brand) terecht komen. De procedure ontleent zijn grote practische betekenis (hij wordt vrijwel algemeen gevolgd) aan zijn eenvoud.

De eerste garage bevat 2 wagens en de eerste brand vraagt 1 wagen. Wij sturen het maximum; in dit geval is dit niet meer dan 1 wagen, dus $x_{11} = 1$. Daarmee is de eerste brand voorzien en behoeven geen wagens vanuit de andere garages naar deze plaats te worden gezonden. Het schema ziet er dan als volgt uit:

	1	3	5
2	1	.	.
3	0	.	.
4	0	.	.

Nu de tweede stap. Wij constateren dat de eerste garage nog een wagen bevat die niet weggereden is. Deze zenden we naar de tweede brand, zodat er een 1 verschijnt onmiddellijk rechts van de 1 van het zojuist gegeven schema. (Wanneer er nog 4 wagens in die garage hadden gestaan, dus meer dan voor de tweede brand noodzakelijk is, dan hadden we het teveel naar de derde brand gestuurd.) Het resultaat is, dat nu niet alleen de eerste kolom sluitend is gemaakt – dit gebeurde al in de eerste stap –, maar ook de eerste rij:

	1	3	5
2	1	1	0
3	0	.	.
4	0	.	.

Bij de volgende stap constateren we, dat voor de tweede brand nog te weinig wagens zijn verzonden: er ontbreken er 2. Deze zijn te vinden in de tweede garage. Het resultaat is dat nu ook de tweede kolom sluitend wordt:

	1	3	5
2	1	1	0
3	0	2	.
4	0	0	.

De vierde stap: zend 1 wagen van de tweede garage naar de derde brand, dan wordt ook de tweede rij sluitend:

	1	3	5
2	1	1	0
3	0	2	1
4	0	0	.

De vijfde en laatste stap: zend de 4 wagens van de derde garage naar de derde brand, dan is het hele systeem sluitend:

	1	3	5
2	1	1	0
3	0	2	1
4	0	0	4

Wij hebben nu inderdaad een oplossing met slechts (precies) vijf positieve getallen. Bij deze procedure merken wij nog het volgende op: (i) Wij kunnen niet in de eerste stap een 2 invullen in de linkerbovenhoek, zodat de eerste *rij* sluitend wordt. Dan staat er n.l. in de eerste kolom een 2 op plaats x_{11}, doch $x_{11} + x_{21} + x_{31} = 1$, en dus zou hétzij x_{21} hetzij x_{31} negatief moeten worden. M.a.w., van de twee, rij of kolom, maken wij die sluitend die het *minste* nodig heeft om sluitend gemaakt te worden.[1] (ii) Omdat er bij iedere stap hétzij een rij, hetzij een kolom sluitend wordt gemaakt, zijn er maximaal 6 positieve getallen. Wij toonden echter reeds aan, dat zodra er aan 5 nevenvoorwaarden is voldaan, de zesde *automatisch* ook klopt. Indien wij dus de vijfde nevenvoorwaarde sluitend maken, dan maken wij automatisch ook de zesde sluitend, en dus zijn er 5 positieve getallen.

Andere voorbeelden:

	4	1	2	10
2	.	.	.	.
6	.	.	.	.
9	.	.	.	.

→

	4	1	2	10
2	2	0	0	0
6	2	1	2	1
9	0	0	0	9

	1	2	8
5	.	.	.
4	.	.	.
1	.	.	.
1	.	.	.

→

	1	2	8
5	1	2	2
4	0	0	4
1	0	0	1
1	0	0	1

In beide voorbeelden is de som der getallen in de linkerkolom gelijk aan de som der getallen in de bovenrij. In beide is het aantal garages plus het aantal branden 7; dus mag het ons niet verwonderen dat er 6 positieve getallen zijn in het antwoord.

Deze linkerbovenhoek – of noordwesthoek – regel is heel eenvoudig, doch maakt *helemaal geen gebruik* van de gegeven kilometerafstanden. We moeten dus i.h.a. verwachten, dat de bereikte oplossing 'te duur' zal zijn. De gegeven oplossing voor

1. Soms hebben rij en kolom *evenveel* nodig om sluitend te worden. Dit gebeurt altijd in de laatste stap, doch soms ook 'halverwege'. Het geeft aanleiding tot minder dan 5 positieve getallen; dit leidt tot enige, zij het geen serieuze complicaties.

het brandweerprobleem, gegeven onder de vijfde stap, leidt tot een totaal aantal kilometers van:

$$5 \times 1 + 2 \times 1 + 8 \times 2 + 6 \times 1 + 7 \times 4 = 57.$$

Wellicht is het mogelijk een begin-oplossing te vinden die minder kilometers vereist. Hiertoe moeten we gebruik maken van het gegeven afstandenschema:

	1	3	5
2	5	2	9
3	3	8	6
4	2	5	7

	1	3	5
2	.	.	.
3	.	.	.
4	.	.	.

Het ligt dan voor de hand met de kortste route te beginnen; of, zo er verscheidene kortste routes zijn, met één daarvan. Dit laatste is hier inderdaad het geval, nl. 2 km van de eerste garage naar de tweede brand en ook van de derde garage naar de eerste brand. Wij kiezen de eerste, $x_{12} = 2$, en maken daarmee de eerste rij sluitend. Vervolgens gaan we steeds wagens verzenden zo dat (i) *die* garage geleegd wordt die bij de vorige stap nog wagens over had resp. naar *die* brand wordt gereden die bij de vorige stap onvoldoende wagens kreeg en (ii) steeds de kortste route wordt gevolgd. Bij de eerste stap kreeg de tweede brand slechts 2 wagens terwijl er 3 nodig zijn; dus zenden wij er 1 heen en wel vanuit de derde garage omdat die het dichtste bij is (5 km tegenover 8 km voor de tweede garage). Daardoor wordt de tweede kolom sluitend. Zo gaan we door met het sluitend maken van de overige rijen en kolommen; het resultaat is na de vijfde stap:

	1	3	5
2	0	2	0
3	0	0	3
4	1	1	2

De totale kilometerafstand die nú moet worden verreden bedraagt

$$2 \times 2 + 6 \times 3 + 2 \times 1 + 5 \times 1 + 7 \times 2 = 43,$$

hetgeen t.o.v. 57 een nogal grote verbetering is. De vraag is nu, of het zo verkregen schema optimaal is. Deze vraag zullen wij thans bezien.

3. WANNEER IS EEN TRANSPORTPLAN OPTIMAAL?

Bij de verificatie van de optimaliteit van een gegeven transportplan wordt gebruik gemaakt van de volgende stelling: de optimale oplossing van een transportprobleem verandert niet, wanneer alle getallen in een rij van het afstandenschema worden verhoogd of verlaagd met eenzelfde getal. Hetzelfde geldt voor alle getallen in een kolom. Nemen we bijv. ons oorspronkelijke afstandenschema,

		1	3	5
I:	2	5	2	9
	3	3	8	6
	4	2	5	7

en trekken we 2 af van de onderste rij:

		1	3	5
II:	2	5	2	9
	3	3	8	6
	4	0	3	5

dan blijft volgens de genoemde stelling de optimale oplossing dezelfde. D.w.z. in I en II dienen we uit elke garage naar elke brand paarsgewijs dezelfde aantallen auto's te zenden. Trekken we vervolgens 4 af van alle getallen in de eerste rij en tellen we 1 op bij alle getallen in de tweede kolom, dan is het resultaat:

		1	3	5
III:	2	1	—1	5
	3	3	9	6
	4	0	4	5

en de optimale oplossing van III is dezelfde als die van I en II. Het lijkt wellicht onzinnig om met negatieve afstanden te werken (van de eerste garage naar de tweede brand in III); en dat is het natuurlijk ook. We moeten echter bedenken dat het hier alleen maar gaat om een wiskundige manipulatie, die voor de verdere afleidingen van belang zal blijken. We zullen nl. zien, dat de tekens van zo'n aangepast afstandenschema (plus, nul, min) van centrale betekenis zijn.

Waarom is het zo, dat de optimale oplossing niet verandert door het optellen of aftrekken van een vast getal bij een rij of een kolom? Om dit te doorzien beschouwen we de doelstellingsfunctie van schema I:

$$5x_{11} + 2x_{12} + 9x_{13}$$
$$+ 3x_{21} + 8x_{22} + 6x_{23}$$
$$+ 2x_{31} + 5x_{32} + 7x_{33}.$$

De doelstellingsfunctie van schema II is precies dezelfde, echter met uitzondering van het feit dat de volgende uitdrukking (betrekking hebbend op de onderste rij) moet worden afgetrokken:

$$2x_{31} + 2x_{32} + 2x_{33} = 2(x_{31} + x_{32} + x_{33}).$$

Maar dit is niets anders dan $2 \times 4 = 8$, omdat de som van de x-en op de onderste rij noodzakelijkerwijs gelijk is aan 4 (er staan immers 4 wagens in de derde garage). Dit betekent, dat de doelstellingsfunctie van II afgeleid kan worden uit die van I door 8 af te trekken. Merk op dat dit onafhankelijk is van de wijze waarop wij de wagens laten rijden; het geldt voor alle denkbare transportplannen, die alle uniform 8 km 'goedkoper' worden. Dan is ook gemakkelijk in te zien dat het optimale transportplan niet verandert: het is in II hetzelfde als in I, alleen de waarde van het minimum wordt met 8 gereduceerd. Voor schema III gaat het analoog. De doelstellingsfunctie wordt uit die van II verkregen door er

$$4(x_{11} + x_{12} + x_{13}) = 4 \times 2 = 8 \quad \text{(eerste rij)}$$

van af te trekken en er vervolgens

$$1(x_{12} + x_{22} + x_{32}) = 1 \times 3 = 3 \quad \text{(tweede kolom)}$$

bij te tellen. Ook hier verandert de optimale oplossing dus niet.

Wij zullen van deze stelling gebruik maken door het afstandenschema zodanig te transformeren, dat voor de routes die inderdaad worden verreden de afstand nul wordt. Vervolgens zal het gaan om de tekens van de overige afstanden (dus van de routes die in het onderzochte transportplan niet worden bereden). Wellicht zal de lezer nu al aanvoelen, dat als die overige afstanden alle positief zijn het transportplan niet te verbeteren is, dus optimaal is; maar wij zullen de procedure stap voor stap behandelen.

Laat ons uitgangspunt zijn het aloude afstandenschema en de op blz. 56 geformuleerde oplossing, verder beginoplossing te noemen:

Afstandenschema

	1	3	5
2	5	2	9
3	3	8	6
4	2	5	7

Beginoplossing

	1	3	5
2	0	2	0
3	0	0	3
4	1	1	2

We gaan nu het afstandenschema bekijken voorzover de routes in de beginoplossing inderdaad worden bereden. Dit betekent dat we de overige afstanden gewoon buiten beschouwing laten. Voorts *randen* wij het schema met een benedenrij en een rechterkolom, in eerste aanleg met stippen:

	1	3	5	
2		2		.
3			6	.
4	2	5	7	.
	.	.	.	

Het gaat er nu om deze gereden afstanden nul te maken door per rij en per kolom zekere getallen op te tellen of af te trekken; hetgeen, naar wij zojuist zagen, niet van invloed is op de optimale oplossing. Daartoe vervangen we de eerste stip op de benedenrij door een in beginsel willekeurig getal, waarvoor men in de practijk 0 pleegt te nemen:

	1	3	5	
2		2		.
3			6	.
4	2	5	7	.
	0	.	.	

en we constateren het volgende: als de afstand 2 van de derde garage naar de eerste brand tot nul gereduceerd moet worden, dan kunnen we dat doen door 2 van de derde rij af te trekken. Dit schrijven we als volgt:

	1	3	5	
2		2		.
3			6	.
4	2	5	7	2
	0	.	.	

hetgeen wordt geïnterpreteerd in de zin, dat de afstand 2 van de derde garage naar de eerste brand wordt opgesplitst in een getal 2 voor de derde rij en een getal 0 voor de eerste kolom. Gegeven de 2 van de derde rij, hebben we een 3 in de tweede kolom nodig om de afstand 5 van de derde garage naar de tweede brand uit te splitsen; en evenzo hebben we een 5 in de derde kolom nodig voor de afstand 7 van de derde garage naar de derde brand. Dus:

	1	3	5	
2		2		.
3			6	.
4	2	5	7	2
	0	3	5	

We zijn er nu bijna; er zijn nog twee stippen. Die van de tweede rij moet 1 worden, want de afstand 6 van de tweede garage naar de derde brand wordt dan uitgesplitst als 1 + 5; en die van de eerste rij —1, zodat de afstand 2 van de eerste garage naar de tweede brand wordt uitgesplitst als —1 + 3. Het resultaat is:

	1	3	5	
2		2		—1
3			6	1
4	2	5	7	2
	0	3	5	

Hiermee is de zaak rond. Elk van de getallen van het inwendige deel van het schema is nu gelijk aan de som van de corresponderende getallen in de benedenrij en de rechterkolom. Gegeven dat het eerste getal in de benedenrij gelijk wordt gesteld aan nul, zijn de overige rij- en kolomgetallen (in een bepaalde volgorde) direct en eenduidig te bepalen. Het is eenvoudig na te gaan dat als wij het eerste getal van de benedenrij gelijk hadden gesteld aan bijv. 3 (hetgeen ook best mag, omdat de 0 immers een willekeurige keuze is), het volgende schema ontstaat:

	1	3	5	
2		2		—4
3			6	—2
4	2	5	7	—1
	3	6	8	

Hier zijn alle getallen van de benedenrij 3 hoger, en alle getallen van de rechterkolom 3 lager. Dit element van keuze, van willekeur, wordt verklaard uit de omstandigheid dat er 6 stippen moeten worden ingevuld aan de hand van slechts 5 getallen in het inwendige van het schema.

Wij keren nu terug naar het volledige afstandenschema, dat we randen met de zojuist gevonden rij- en kolomgetallen:

	1	3	5	
2	5	2	9	—1
3	3	8	6	1
4	2	5	7	2
	0	3	5	

Vervolgens gaan we van alle afstanden de bijbehorende rij- en kolomgetallen aftrekken. Dus wordt bijv. de afstand 8 van de tweede garage naar de tweede brand verminderd met 3 en 1, wordt dus 4. Vollediger:

	1	3	5
2	5—(—1)—0 = 6	2—(—1)—3 = 0	9—(—1)—5 = 5
3	3—1—0 = 2	8—1—3 = 4	6—1—5 = 0
4	2—2—0 = 0	5—2—3 = 0	7—2—5 = 0

en korter:

	1	3	5
2	6	0	5
3	2	4	0
4	0	0	0

Dit afstandenschema heeft nullen voor alle routes die in de beginoplossing worden bereden; hetgeen ons niet mag verwonderen, want onze rij- en kolomgetallen waren juist voor dat doel zo geconstrueerd! Bovendien heeft het transportprobleem behorend bij dit afstandenschema dezelfde optimale oplossing als dat behorend bij de oorspronkelijke afstandentabel; dat volgt uit onze stelling inzake het niet veranderen van die oplossing in het geval van optellen of aftrekken van vaste getallen per rij of kolom. Maar het nieuwe, getransformeerde transportprobleem heeft een bijzonder eenvoudige oplossing!

Om dit te doorzien beschouwen we zijn doelstellingsfunctie:

$$6x_{11} + 0x_{12} + 5x_{13}$$
$$+ 2x_{21} + 4x_{22} + 0x_{23}$$
$$+ 0x_{31} + 0x_{32} + 0x_{33}.$$

De minimale waarde die deze doelstellingsfunctie kan aannemen is nul. Een negatieve waarde is nl. onmogelijk vanwege de niet-negativiteitsrestricties van de x-en, die hier elk vermenigvuldigd worden met een coëfficiënt die positief of nul is. De doelstellingsfunctie kan wèl de waarde nul aannemen. Daartoe moet men ervoor zorgen, dat de x-en met positieve coëfficiënten (x_{11}, x_{13}, x_{21}, x_{22}) alle nul zijn. Maar onze beginoplossing doet nu juist dit! Hij luidt:

$$x_{12} = 2, \quad x_{23} = 3, \quad x_{31} = 1, \quad x_{32} = 1, \quad x_{33} = 2$$

en alle overige x-en zijn nul. En dat zijn nu juist de x-en waarvoor we moeten zorgen dat ze nul zijn! Onze beginoplossing leidt dus tot de minimale waarde (nl. nul) van de doelstellingsfunctie bij het getransformeerde transportprobleem. Hij is dus de optimale oplossing van dat probleem en dus ook van het oorspronkelijke transportprobleem waarvoor we ons in werkelijkheid interesseren.

Daarmee hebben wij de oplossing gevonden en we weten dus nu bovendien, dat 43 kilometer het minimale kilometertotaal is. We hebben ook gevonden hoe men kan verifiëren of een bepaald transportplan inderdaad de optimale oplossing van het transportprobleem is. Men transformeert het afstandenschema zodanig, dat de bereden routes 'afstand nul' krijgen. Zijn dan de niet bereden afstanden na transformatie alle positief, dan hebben we de oplossing die we zoeken te pakken. Zijn er negatieve waarden bij, dan moeten we verder zoeken. Hoe – dat zullen we nu zien.

4. HOE MAKEN WE EEN NIET-OPTIMAAL TRANSPORTPLAN OPTIMAAL?

Om misverstand te voorkomen vermelden we uitdrukkelijk, dat de wijze waarop wij onze beginoplossing hebben opgesteld niet noodzakelijkerwijs onmiddellijk leidt tot het optimum. Dit kan zo zijn, maar het is lang niet altijd het geval. Dan dienen we de beginoplossing successievelijk te verbeteren, stap voor stap, tot we uiteindelijk in het optimum terecht

komen. Om dit te laten zien zullen we nu als beginoplossing het transportplan van de noordwesthoekregel nemen (zie blz. 54):

Afstandenschema

	1	3	5
2	5	2	9
3	3	8	6
4	2	5	7

Beginoplossing

	1	3	5
2	1	1	0
3	0	2	1
4	0	0	4

De eerste stap bestaat uit het nagaan of we de optimale oplossing al hebben (in het onderhavige geval strikt genomen overbodig, omdat we al een andere oplossing met een geringer kilometertotaal hebben gevonden!). We nemen dus het afstandenschema, laten die routes weg die door de oplossing in kwestie niet worden bereden; en berekenen de rij- en kolomgetallen voor de uitsplitsing van de resterende afstanden. Het resultaat is:

	1	3	5	
2	5	2		5
3		8	6	11
4			7	12
	0	—3	—5	

Vervolgens trekken we van elke afstand de bijbehorende rij- en kolomgetallen af. Het resultaat is:

	1	3	5
2	5— 5—0 = 0	2— 5—(—3) = 0	9— 5—(—5) = 9
3	3—11—0 = —8	8—11—(—3) = 0	6—11—(—5) = 0
4	2—12—0 = —10	5—12—(—3) = —4	7—12—(—5) = 0

of korter:

	1	3	5
2	0	0	9
3	—8	0	0
4	—10	—4	0

Het nieuwe schema heeft ‘negatieve afstanden’, drie stuks in totaal. De bijbehorende doelstellingsfunctie is

$$0x_{11} + 0x_{12} + 9x_{13}$$
$$- 8x_{21} + 0x_{22} + 0x_{23}$$
$$- 10x_{31} - 4x_{32} + 0x_{33}.$$

Kennelijk kunnen we de waarde van de doelstellingsfunctie kleiner maken (ons doel is minimeren!) door x_{21} of x_{31} of x_{32} positief te maken. (In de beginoplossing zijn ze alle nul.) De eerste vraag is: welk van de drie zullen we kiezen? We nemen x_{31}, omdat die met het meest negatieve getal (—10) wordt vermenigvuldigd. De tweede vraag: hoe groot zullen we x_{31} maken? Antwoord: zo groot mogelijk (iedere extra vergroting leidt tot een reductie van de waarde van de doelstellingsfunctie), maar we moeten de nevenvoorwaarden in het oog houden.

We beschouwen opnieuw de beginoplossing en plaatsen een plus bij de 0 van x_{31} om aan te geven, dat we x_{31} tot een positieve waarde gaan ophogen:

	1	3	5
2	1	1	0
3	0	2	1
4	0^+	0	4

Gaan we één of meer wagens versturen van de derde garage naar de eerste brand, dan heeft dat twee onmiddellijke gevolgen. In de eerste plaats dienen er dan minder wagens vanuit andere garages naar de eerste brand te gaan; in de tweede plaats dienen er ook minder vanuit de derde garage naar de andere branden te gaan. Dit leidt tot de volgende opstelling:

	1	3	5
2	1^-	1	0
3	0	2	1
4	0^+	0	4^-

Er zijn twee mintekens toegevoegd, een linksboven en een rechtsonder. Zij geven aan, dat deze getallen gereduceerd moeten worden, hetgeen trouwens rechtstreeks uit de zojuist gegeven argumentatie volgt. Uiteraard kunnen we het aantal wagens vanuit de tweede garage naar de eerste brand en die vanuit de derde garage naar de tweede brand niet reduceren, want die aantallen zijn al nul.

We zijn er echter nog niet. Als we nl. x_{11} reduceren, dus

minder wagens uit de eerste garage naar de eerste brand zenden, dan moeten er meer uit die garage naar andere branden. Evenzo, als we x_{33} reduceren omdat de derde garage minder te leveren heeft (doordat x_{31} tot een positieve waarde wordt opgehoogd), dan zijn extra wagens nodig voor de derde brand. Aldus krijgen we het volgende schema:

	1	3	5
2	1^-	1^+	0
3	0	2	1^+
4	0^+	0	4^-

(We hadden natuurlijk ook de 0 rechtsboven van een plus kunnen voorzien, maar we hebben al zo'n 0 met een plus, nl. linksonder; de procedure komt erop neer dat we ons tot één zo'n exemplaar beperken en verder onze krachten op de positieve getallen beproeven.) Zo komen we tot een 'rondedans van plussen en minnen', die bijna volledig is. We moeten alleen nog rekening houden met het feit, dat meer wagens van de eerste garage naar de tweede brand (de plus van $x_{12} = 1$) een reductie van het aantal wagens van andere garages naar die brand noodzakelijk maakt; evenzo dat meer wagens van de tweede garage naar de derde brand (de plus van $x_{23} = 1$) leidt tot een reductie van het aantal wagens van die garage naar andere branden. In beide gevallen leidt dit tot een min bij het centrale element ($x_{22} = 2$):

	1	3	5
2	1^-	1^+	0
3	0	2^-	1^+
4	0^+	0	4^-

Hiermee is de rondedans compleet. We zijn begonnen met een nul met een plus en nu, aan het eind, hebben we één plus en één min in elke rij en in elke kolom. Tot hoever kunnen we nu x_{31} per saldo ophogen? Daartoe dienen we te bedenken, dat alleen de minnen gevaarlijk zijn. Met iedere wagen waarmee we x_{31} vanaf nul ophogen reduceren we het aantal over de met min aangegeven routes, en wel ook met één wagen. We kunnen dus kennelijk maximaal gaan tot het *laagste* getal dat van een min voorzien is, anders zou dit getal negatief worden.

Hier is dit laagste getal 1 (zie linksboven), dus verhogen we x_{31} van 0 tot 1. Alle routes met een plus worden dan ook met 1 verhoogd, alle met een min met 1 verlaagd, en het resultaat is:

	1	3	5
2	0	2	0
3	0	1	2
4	1	0	3

en het kilometertotaal is nu

$$2 \times 2 + 8 \times 1 + 6 \times 2 + 2 \times 1 + 7 \times 3 = 47,$$

dus een verbetering van 10 km t.o.v. de beginoplossing.[1]

Hebben we het optimum bereikt? Neen. Want het afgekorte afstandenschema met bijbehorende rij- en kolomgetallen is nu:

	1	3	5	
2		2		—5
3		8	6	1
4	2		7	2
	0	7	5	

en het afstandenschema na aftrek van die rij- en kolomgetallen is dan

	1	3	5
2	5—(—5)—0 = 10	2—(—5)—7 = 0	9—(—5)—5 = 9
3	3—1—0 = 2	8—1—7 = 0	6—1—5 = 0
4	2—2—0 = 0	5—2—7 = —4	7—2—5 = 0

zodat er nog een negatief getal staat, nl. —4 op de plaats van x_{32}. We moeten dus x_{32} laten toenemen:

	1	3	5
2	0	2	0
3	0	1	2
4	1	0^{+}	3

1. Het zal nu ook duidelijk zijn waarom we slechts één 0 van een plus hebben willen voorzien. Waren het er nl. twee geweest, dan waren er twee getallen in het schema positief geworden die eerst 0 waren. Hiertegenover staat slechts één positief getal dat 0 geworden is. Daarmee zou het aantal positieve getallen 6 geworden zijn i.p.v. 5.

Dit geeft aanleiding tot de volgende rondedans:

	1	3	5
2	0	2	0
3	0	1^-	2^+
4	1	0^+	3^-

(Het getal 1 in de linkerbenedenhoek kan geen min krijgen, omdat er dan in de eerste kolom ook een positief getal een plus zou moeten krijgen; doch er is geen positief getal meer over in de eerste kolom.) Wij kunnen x_{32} maximaal 1 laten worden omdat het laagste getal met een min een 1 is. Zo vinden wij

	1	3	5
2	0	2	0
3	0	0	3
4	1	1	2

Dat is de optimale oplossing, zoals al bekend!

5. EEN QUIZ

Wij zullen nu een toepassing laten zien die op het eerste gezicht helemaal niets met transportvraagstukken te maken heeft. Een quiz: vier deelnemers in een team, die ieder vragen 'op hun gebied' beantwoorden. De vier onderwerpen waarover wordt gevraagd zijn Aardrijkskunde (*A*), Geschiedenis (*G*), Muziek (*M*) en Sport (*S*). De vier deelnemers zijn de heren Jansen (*J*), Klaassen (*K*), Pietersen (*P*) en Willemsen (*W*). Ieder van de vier heeft op alle gebieden een zekere bekwaamheid, maar sommigen blinken uit in speciale onderdelen. Wij nemen nu aan dat deze bekwaamheid evenals op school wordt aangeduid met een cijfer, variërend van nul tot tien. De bekwaamheden van de heren *J*, *K*, *P* en *W* kunnen nu als volgt worden samengevat in tabelvorm:

	J	*K*	*P*	*W*
A	8	8	4	9
G	6	8	5	7
M	5	8	5	8
S	8	9	7	4

Zoals uit deze tabel blijkt, is de heer Klaassen de bolleboos van het gezelschap, terwijl de heer Pietersen wat zwakjes staat.

Ieder van de vier heren moet vragen beantwoorden op één en slechts één gebied. Het probleem is wie vragen op welk gebied moet beantwoorden om te verzekeren dat het team als geheel zoveel mogelijk presteert. Uiteraard moet dit laatste nader worden gepreciseerd. Daartoe gaan we ervan uit, dat als Jansen vragen op het terrein van Geschiedenis beantwoordt, zijn prestatie samenvalt met zijn cijfer (dus een 6), en evenzo voor de overige personen en terreinen, en dat de prestatie van het team als geheel gemeten wordt door de som van de vier individuele scores. We nummeren vakken en personen beide van 1 tot en met 4 (in de volgorde van de cijfertabel) en schrijven $6x_{21}$ voor de score die voor Geschiedenis (2) door Jansen (1) wordt geboekt. Neemt Jansen inderdaad Geschiedenis voor zijn rekening, dan is $x_{21} = 1$; doet hij het niet, dan is $x_{21} = 0$. Enzovoorts. De doelstellingsfunctie wordt dus:

$$\begin{aligned} &8x_{11} + 8x_{12} + 4x_{13} + 9x_{14} \\ &+ 6x_{21} + 8x_{22} + 5x_{23} + 7x_{24} \\ &+ 5x_{31} + 8x_{32} + 5x_{33} + 8x_{34} \\ &+ 8x_{41} + 9x_{42} + 7x_{43} + 4x_{44}. \end{aligned}$$

Dus een lineaire doelstellingsfunctie in 16 beslissingsvariabelen, die geen van alle negatief mogen zijn. Nu de nevenvoorwaarden. Die komen hierop neer, dat elke persoon gaat over één terrein, niet meer en niet minder. Een voorbeeld:

		J	*K*	*P*	*W*
		1	1	1	1
A	1	1	0	0	0
G	1	0	1	0	0
M	1	0	0	0	1
S	1	0	0	1	0

In dit voorbeeld wordt de Aardrijkskunde aan Jansen toevertrouwd, de Geschiedenis aan Klaassen, de Muziek aan Willemsen en de Sport aan Pietersen. Men ziet dat de structuur van dit vraagstuk in volmaakte overeenstemming is met dat van een gewoon transportprobleem. Als we *A*, *G*, *M* en *S* als vier garages met elk één brandweerauto zouden beschouwen, en *J*, *K*, *P* en *W* als vier branden (licht-alarm), dan zou er ook overeenstemming qua inhoud zijn.

Het hier gegeven voorbeeld is tevens de optimale oplossing; de totaalscore bedraagt $8 + 8 + 8 + 7 = 31$, het maximaal bereikbare. Het zou ons te ver voeren de afleiding te geven. Het aantal nevenvoorwaarden is $4 + 4 — 1 = 7$, zodat men zou verwachten dat dit ook het aantal positieve variabelen in de oplossing is. In feite moet dit laatste aantal noodzakelijkerwijs 4 zijn, zoals men gemakkelijk zal inzien. Dit is van enige invloed op de berekeningen.

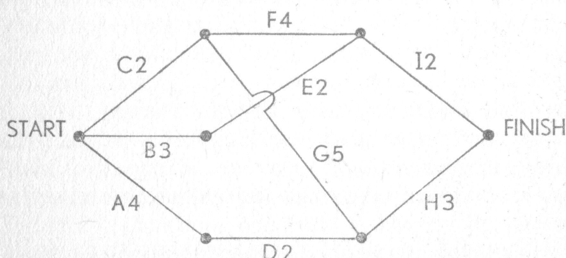
6. DE GEDACHTENGANG VAN HET KRITIEKE PAD

Hebben wij zojuist een (quiz) probleem bezien, waarbij transport, routes en paden op het eerste gezicht volmaakt irrelevante begrippen waren, nu zullen wij het accent verleggen naar een terrein met paden, waar niet het kortste maar het *langste* pad in het centrum van de belangstelling staat. Hierbij verlaten wij het terrein van de lineaire programmering. In plaats van transportproblemen zullen wij meer algemene organisatieproblemen bezien. 'Organisatie' is hier gebruikt in een veelomvattende betekenis. Enerzijds kan men hierbij denken aan een eenvoudige dagindeling (thee zetten, afwassen, stof afnemen,...) of een minder alledaagse dagindeling (verjaarspartijtje organiseren: uitnodigingen verzenden, inkopen doen, naar kapper gaan,...). Deze taken zijn duidelijk te overzien; de organisatie die zij vereisen is ófwel een routinezaak, ófwel in één avond te regelen met potlood, papier, telefoon en echtgenoot. Anderzijds kan men denken aan de organisatie benodigd voor het bouwen van een huis, fabriek, zeeschip of Euromast [aanbesteden, heien, materialen bestellen (levertijden!), beton storten, leidingen aanleggen, restaurant bouwen, restaurant omhoog hijsen]. Dit soort projecten vereist een ver doorgevoerde planning. Het doel is hierbij zo mogelijk knelpunten ('bottlenecks') te vermijden; of in ieder geval te weten te komen waar zich mogelijke knelpunten kunnen voordoen. Dit alles ten behoeve van een zo snel en efficiënt mogelijke bouw.

De methode die wij zullen bespreken is een (vereenvoudigde) versie van de in Amerika ontwikkelde *Critical Path Method* (CPM). Dit is in eerste aanleg hetzelfde als de zgn. PERT, hetgeen de eerste-letterafkorting is van *Program Evaluation and Review Technique*. Ook PERT werd in Amerika ontwikkeld,

ongeveer gelijktijdig (1956-1957) met CPM. Men zegt dat dankzij dit nieuwe hulpmiddel bij de organisatie de bouw van de eerste Polaris-raket onderzeeboot twee jaar eerder klaar was dan aanvankelijk was geschat. Voor ieder, die weet dat dergelijke grote aanbestedingen nogal eens jaren te laat worden opgeleverd, zal deze prestatie nog meer aanspreken.

Zowel bij CPM als bij PERT bestaat de eerste stap uit het analyseren van de verschillende deeltaken en het bepalen hoeveel tijd ieder van deze deeltaken zal vereisen. Deze gegevens worden dan gebruikt bij het construeren van een netwerk; de knooppunten van dit netwerk worden gevormd door de (voltooide) deeltaken. *Die* knooppunten in het netwerk zijn verbonden, waarvan het *noodzakelijk* is dat zij *na elkaar* geschieden. Andere deeltaken kunnen *gelijktijdig* worden aangepakt. Beschouwen wij nu – in abstracto – een project dat bestaat uit 9 deeltaken *A*, *B*, *C*, . . ., *I* en dat gerepresenteerd kan worden door het volgende netwerk:

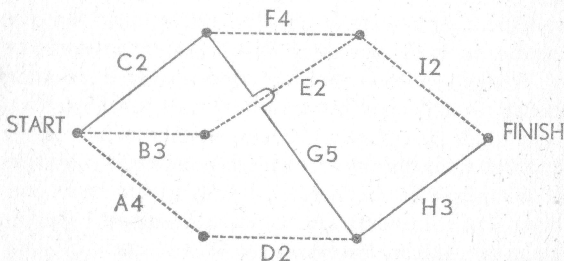


Uit dit netwerk blijkt dat de voltooiïng van *A* 4 dagen kost; en van deeltaak *H* 3 dagen. Oók blijkt eruit dat taak *C* klaar moet zijn vóór met *F* of *G* kan worden begonnen. De taken *F* en *G* kunnen echter (evenals *A*, *B* en *C*) gelijktijdig worden uitgevoerd. Voor men met taak *H* kan beginnen moeten zowel taak *D* als taak *G* zijn voltooid. Enzovoorts.

Wij wensen nu het *kritieke pad* van dit netwerk te bepalen. Dit kritieke pad wordt gevormd door een reeks van taken zodanig dat, indien zich in een dier taken een vertraging zou voordoen, de voltooiïng van het hele project zou worden vertraagd. En omgekeerd: indien een der taken op het kritieke pad sneller is voltooid dan men aanvankelijk had geschat, dan is het hele project eerder voltooid. Het kritieke pad in een der-

gelijk netwerk is *het langste pad* (in dagen gemeten) om van het beginpunt tot het eindpunt te komen. Uit het netwerk blijkt bijv. dat het minimaal 7 dagen duurt voor men met taak *H* kan beginnen. Weliswaar heeft men voor *A* en *D* samen slechts 6 dagen nodig, doch de weg *CG* vereist 7 dagen (en men kan met *H* niet aanvangen vóór *G* is voltooid).

Een en ander betekent dat de taken *A* en *D niet kritiek* zijn; enige vertraging is mogelijk zonder dat dit de totale tijdsduur nodig om het project af te maken zal beïnvloeden. Doch de taken *C* en *G* zijn *wél* kritiek. Het kritieke pad is hieronder weergegeven door de niet-kritieke taken gestippeld te tekenen; in dit eenvoudige geval is het door observatie bepaald. In gecompliceerde gevallen kan het pad slechts met een electronische rekenmachine worden bepaald. Het is dan echter zeer eenvoudig, en gaat zeer snel.[1]



Men ziet nu met één oogopslag waar zich knelpunten kunnen voordoen; te weten bij de voltooiïng van *C*, *G* of *H*. Als *A*, *B*, *D*, *E*, *F* of *I* iets langer zouden duren dan geschat, dan beinvloedt dit de tijd benodigd voor het hele project niet. Anders wordt dit, wanneer ze véél langer gaan duren. Bijv.:

Als *A* 6 dagen zou duren, wordt het kritieke pad *ADH*.
Als *I* 5 dagen zou duren, wordt het kritieke pad *CFI*.
Als *E* 5 dagen zou duren zijn *BEI* én *CGH* beide kritiek.
Als *G* slechts 2 dagen zou duren, dan is *ADH* het kritieke pad.

1. Er is bovendien een zgn. Pert-o-graph in de handel, die in Amerika slechts 1 dollar kost. Dit is een discus-vormig zakrekenmachientje, te verglijken met een rekenliniaal, dat speciaal voor het oplossen van kritieke-padproblemen is geconstrueerd. Het kan zeer nuttige diensten bewijzen voor niet ál te gecompliceerde netwerken.

Als *G* slechts 4 dagen zou duren, zijn *CGH* en *ADH* beide kritiek.

Het meest voor de hand liggende gebruik dat van de kritiekepadmethode wordt gemaakt is het inzetten van extra mankracht of machines, of het in overtijd werken *voor die deeltaken* welke op het kritieke pad liggen. De besparingen die hieruit kunnen voortvloeien wegen vaak ruimschoots op tegen de extra kosten. Wat dit betreft heeft deze methode gedurende zijn nog geen tienjarig bestaan zijn sporen reeds verdiend. Dit inzetten van extra mankracht op de kritieke deeltaken gaat vaak ten koste van de mankracht gebruikt bij andere (nietkritieke) deeltaken, die daardoor iets langer gaan duren. Dit voorkomt lange wachttijden en leegloop aan de ene kant, en spanningen en haast aan de andere kant.

Bovendien is het schema bij uitstek goed leesbaar en begrijpelijk. De opzichter kan zich doorlopend eenvoudig oriënteren. *En hij kan het netwerk aanpassen* als de laatste gegevens betreffende vertragingen e.d. bekend worden. Op eenvoudige wijze kan dan voortdurend het verschijnen of verdwijnen van nieuwe resp. oude knelpunten worden bijgehouden. De verschillende afdelingen die bij het project betrokken zijn zullen elkaars problemen beter begrijpen doordat zij een beter overzicht hebben. Dit zal onderlinge fricties en rivaliteit beperken, een niet te onderschatten voordeel.

7. HET KRITIEKE PAD VAN DIT BOEK

Stel U voor dat U overweegt met drie man (*X*, *Y* en *Z*) een boek te schrijven over kwantitatieve economie en operationele research. (Deze situatie deed zich voor toen de uitgever half januari 1963 het Econometrisch Instituut benaderde.) De eerste vraag is of dit plan moet worden uitgevoerd. De tweede: zo ja, hoe, wat zijn de onderwerpen en de hoofdstukken, wie schrijft welke hoofdstukken in welke volgorde, enz. Rekening moet worden gehouden met diverse andere verplichtingen, die voor de ene auteur zwaarder kunnen wegen dan voor de andere: college geven, tentamen afnemen, lopende eigen onderzoekingen, leiding aan onderzoekingen van anderen, conferenties e.d. De eerste ruwe versie diende medio juli klaar te zijn opdat

de afwerking in de rest van de zomer zou kunnen plaats hebben. De zomer is nu eenmaal wat rustiger in het academisch bedrijf, dan zijn er geen colleges en weinig studenten en de meeste buitenlandse onderzoekers zeggen het Instituut vaarwel om zonniger streken op te zoeken. Bovendien is er dan meer typetijd beschikbaar.

Om de ruwe versie gaat het hier. Hieronder volgt een lijst van hoofdstukken zoals die er oorspronkelijk uit zag, met voor elk hoofdstuknummer de naam van de schrijver. (Op zichzelf is het toekennen van de hoofdstukken aan de diverse auteurs een probleem verwant aan dat van de quiz, maar zover zijn we niet gegaan.) Twee hoofdstukken zijn gesplitst (7a, 7b, 13a, 13b) i.v.m. het feit dat die onderdelen door verschillende auteurs zouden worden geschreven.

X	1.	Lineair programmeren
X	2.	Het optimale en het kritieke pad
X	3.	Aan- en afvoeranalyse
Z	4.	Macro-economische modellen
Y	5.	Economische voorspellingen
Y	6.	Economische weerberichten
Y	7a.	Onzekerheid, waarschijnlijkheid en simulatie
X	7b.	Idem
Y	8.	Wiskundige statistiek
Z	9.	Om de gulden van de consument
Z	10.	Wachten
Y	11.	Het strategiebegrip
X	12.	Speltheorie en beleidsspelen
Z	13a.	Productie- en voorraadbeslissingen
Y	13b.	Idem

Wat de benodigde tijd per hoofdstuk betreft, gegeven de andere verplichtingen komt deze neer op 4 weken per hoofdstuk voor *X*, 3 weken voor *Y* en 6 weken voor *Z*. Dit betreft de volledige hoofdstukken, dus exclusief 7 en 13. De tijd benodigd voor 7a, 7b, 13a, en 13b is resp. 3, 1, 1 en 3 weken.

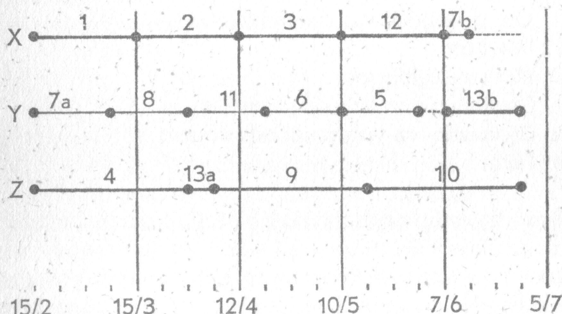
Tenslotte hebben we nog rekening te houden met de volgorde van de hoofdstukken, zulks om drie redenen. Ten eerste schrijft geen auteur twee of meer hoofdstukken tegelijkertijd. Ten tweede is er de samenhang van de verschillende hoofdstukken. In de derde plaats verdient het de voorkeur om de

hoofdstukken zoveel mogelijk in de natuurlijke volgorde (1, 2, . . ., 13) af te doen. Dit leidt tot de volgende regels:

1, 2 en 3 moeten in die volgorde geschreven worden;
3 gaat vooraf aan 7b en 12;
5 kan pas geschreven worden als 3 en 4 klaar zijn;
7a gaat vooraf aan 7b, 8, 10 en 11;
8 gaat vooraf aan 9;
11 gaat vooraf aan 12 en 13b;
13a gaat vooraf aan 13b.

Bovendien is de gelijktijdigheid van het schrijven van de volgende hoofdstukken uitgesloten (i.v.m. het feit dat ze door dezelfde auteur worden geschreven): 5, 6 en 7a; 7b en 12; 4, 9 en 10. De overige ongewenste mogelijkheden van gelijktijdigheid zijn al geëlimineerd door de zeven hierboven opgesomde regels.

Er is met verschillende schema's geëxperimenteerd, alle op basis van de veronderstelling dat op 15 februari met het schrijven zou worden begonnen. Het meest overzichtelijke schema was er een weliswaar zonder knooppunten, maar toch duidelijk van het kritieke-padtype zoals hieronder zal blijken. Het ziet er als volgt uit:



Het hele pad van *Z* is kritiek; bovendien is het pad 1-2-3-12-13b kritiek. Beide zijn daarom met dikke lijnen aangegeven. *Y* heeft in totaal voor de Hoofdstukken 7a, 8, 11, 6 en 5 één week extra. Hij kan desnoods de volgorde van Hoofdstukken 11 en 6 omdraaien, maar dat zou als nadeel hebben dat bij vertraging het

werk aan Hoofdstuk 12 zou moeten worden opgeschoven; en 12 maakt deel uit van een kritiek pad. Voorts moet *Y* beginnen met 7a en 8, omdat anders *Z* niet met 9 kan beginnen (en het is plezierig, hoewel niet strikt noodzakelijk, 9 en 10 in die volgorde te schrijven). Verder moet *Y* met het schrijven van 5 wachten tot 3 (en 4) klaar zijn. Het is wel aardig dat uit deze analyse volgt, dat het eerste deel van Hoofdstuk 7, dus 7a, al heel in het begin moet worden geschreven, doch dat het onderdeel 7b pas op het eind hieraan moet worden toegevoegd. Iets soortgelijks geldt voor de onderdelen 13a en 13b van Hoofdstuk 13. Hoewel *X* op het eind twee weken over heeft, vormt hij toch een belangrijk deel van het kritieke pad in het begin.

Het is opvallend hoe gemakkelijk uit zo'n schema kan worden afgelezen wie wat wanneer zal doen. Bovendien werd duidelijk dat het zelfs met enige tegenslag mogelijk moest zijn een en ander vóór half juli te voltooien – althans in eerste ronde. Zo ving het werk inderdaad 15 februari aan. En vandaag is het, zoals U aan 't schema kunt zien, bijna twee maanden later.

LITERATUUR

Zowel de formulering als de oplossingstechniek van transportproblemen is voorafgegaan aan die van meer algemene lineaire programmeringsproblemen. Het probleem werd het eerst besproken door Hitchcock [1] en daarna in detail uitgewerkt door onze vroegere landgenoot Koopmans [2]. Het transportprobleem wordt uitvoerig besproken in de boeken van Gass en Vajda, die aan het eind van Hoofdstuk 1 vermeld werden. Overigens zijn er ook andere technieken dan de hier behandelde. De in dit hoofdstuk beschreven oplossingstechniek is zodanig, dat doorlopend aan alle nevenvoorwaarden is voldaan; stap voor stap wordt een betere oplossing gevonden, waarbij de in totaal verreden afstand geringer is en nog steeds aan alle nevenvoorwaarden voldaan is. Het kan ook anders. Ford en Fulkerson [3] beschrijven een methode, waarbij niet in alle successieve stappen voldaan is aan alle nevenvoorwaarden, maar waarbij het verreden kilometertotaal geminimeerd wordt, gegeven de nevenvoorwaarden waaraan wèl is voldaan. Men werkt dan successievelijk toe naar het voldoen aan àlle nevenvoorwaarden.

[1] Hitchcock, F. L., 'Distribution of a Product from Several Sources to Numerous Localities'. *Journal of Mathematical Physics*, Vol. 20 (1941), pp. 224-230.

[2] Koopmans, T. C., 'Optimum Utilization of the Transport-

ation System'. *Econometrica*, Supplement to Vol. 17 (1949), pp. 136-145.

[3] Ford, L. R., and D. R. Fulkerson, 'Solving the Transportation Problem'. *Management Science*, Vol. 3 (1956), pp. 24-32.

Het quiz-probleem wordt besproken door Kuhn [4] en in het Nederlands door de Tilburgse lector Lips [5]. Ook dit soort problemen heeft zijn 'eigen' oplossingstechniek.

[4] Kuhn, H. W., 'The Hungarian Method for the Assignment Problem'. *Naval Research Logistics Quarterly*, Vol. 2 (1956), pp. 83-97.

[5] Lips, L., *Wiskunde en economie*. Openbare Les. P.Noordhoff N.V., Groningen. 1961.

Een nuttig boek voor de wiskundige leek op het terrein van het kritieke pad is dat van Stilian e.a. [6]. Een korte leerzame inleiding is [7]. In het Nederlands is het artikel van Reith [8] beschikbaar, dat een aardig overzicht geeft van PERT- en CPM-technieken onder allerlei verschillende omstandigheden. Lezers die niet wiskundig geschoold zijn kunnen alle paragrafen met wiskunde overslaan; er blijft voldoende over dat de moeite van het lezen waard is.

[6] Stilian, G. N., e.a. *PERT, A New Management Planning and Control Technique*. American Management Association, Inc., New York. 1962.

[7] Boehm, G. A. W., 'Helping the Executive to Make up his Mind'. *Fortune*, April 1962.

[8] Reith, P. F., 'Project-planning methoden'. *Statistica Neerlandica*, Jaargang 17 (1963), pp. 137–173.

3. AAN- EN AFVOERANALYSE

1. SAMENHANG IN DE ECONOMIE

Wij staan er nauwelijks bij stil, omdat wij er zo aan gewend zijn: onze gecompliceerde samenleving, waar alles op elkaar moet zijn afgestemd en op elkaar moet aansluiten, werkt geruislozer dan de treinenloop. En zulks zonder spoorboekje. Er zijn agenten, verpleegsters en onderwijzers, hoewel niemand wordt voorgeschreven agent, verpleegster of onderwijzer te worden. Er is staal en ijzer voor de buizenfabrieken, er zijn hoogovens voor dat staal en dat ijzer, arbeiders voor die hoogovens, eieren voor die arbeiders, kippen voor die eieren, boeren voor die kippen – en transportbedrijven om al deze producten te vervoeren van de ene producent naar de andere, naar de handelaars, naar de consument. De bekendste grondlegger van de economische wetenschap – de Schot Adam Smith uit de tweede helft van de achttiende eeuw – sprak in dit verband van een 'onzichtbare hand' die dit alles regelt en organiseert. Die hand is sinds de achttiende eeuw, en in het bijzonder sinds de Tweede Wereldoorlog, misschien iets zichtbaarder geworden, omdat de herinnering aan de depressie van de dertiger jaren ons ervan overtuigd heeft dat de geruisloze werking toch niet in alle opzichten ideaal is. Menen wij dus een zichtbare hand te onderkennen, dan is dat die van de overheid; maar goed, de samenleving is sinds de achttiende eeuw ook wat gecompliceerder geworden. Er blijft voldoende reden tot verwondering.

De 'verklaring' van dit alles met een 'onzichtbare hand' doet een beetje metafysisch aan. Het verdient aanbeveling wat nader op de zaak in te gaan en in eerste instantie zullen we dit meer 'beschrijvend' dan 'verklarend' doen. We zullen ons bezig houden met transacties van het ene bedrijf met het andere, of liever: van een aantal groepen van bedrijven onderling.[1] Men kan dit beschouwen als een soort van economisch transportstelsel, en in dit opzicht sluit het huidige hoofdstuk dan ook aan op zijn voorganger. Ieder bedrijf produceert, draagt het geproduceerde over aan een ander bedrijf, dat er op zijn beurt weer

1. Soortgelijk werk is ook wel verricht voor de onderlinge relaties van afdelingen binnen een enkel bedrijf, maar de ontwikkeling op dit 'micro-economische' terrein is veel bescheidener.

iets mee doet, daarna gaat het naar een derde bedrijf, enz., tot het uiteindelijk bij de consument terecht komt. De 'vader' van het type van onderzoek, dat in dit hoofdstuk ter sprake komt, is Wassily W. Leontief van Harvard University. Zijn standaardwerk op dit terrein werd in 1941 gepubliceerd.

De eerste stap houdt in, dat we het geheel van alle bedrijven van een land ('het bedrijfsleven') in een aantal sectoren opdelen. Er zijn gedetailleerde schema's waarbij men 250 verschillende sectoren onderscheidt; en ook wat meer geaggregeerde (d.w.z. samengevoegde) schema's van niet meer dan 30 of 40 sectoren. Maar de beginselen kunnen even goed worden geïllustreerd aan de hand van een voorbeeld waarbij men slechts drie sectoren onderscheidt, bijv. de sector landbouw, de sector industrie en de sector diensten. Deze laatste sector omvat dan o.a. het bank- en verzekeringswezen, de handel, de kappers, de horecabedrijven en de transportdiensten. Iedere activiteit moet in één en precies één dezer sectoren worden ondergebracht. Hierover mag geen misverstand bestaan. Er moet dus worden vastgesteld of bijv. de zuivelfabrieken tot de sector landbouw of de sector industrie gerekend dienen te worden. (Men pleegt ze tot de sector industrie te rekenen.) Het is bepaald niet zonder consequenties welke oplossing men kiest; hier zullen wij echter aannemen, dat deze kwesties zijn geregeld.

De volgende stap is het vaststellen van een *transactieschema*, waaruit blijkt hoeveel iedere sector in een zeker jaar aan iedere sector verkoopt. Deze bepaling is in de praktijk een vak op zichzelf, maar in theorie betrekkelijk eenvoudig. Als een boer een maaimachine koopt geldt dat als een levering van de sector industrie aan de sector landbouw. Als een buizenfabriek ijzer koopt geldt dat als een levering *van* de sector industrie *aan* de sector industrie. Vanzelfsprekend echter wordt niet alles wat in de landbouw, industrie of dienstensector wordt *geproduceerd* ook in een dezer sectoren *verwerkt*. Een belangrijk deel van de productie zal in de regel gaan naar *de consument*, die van de sector landbouw zijn eieren, van de industrie zijn bromfiets, en van de dienstensector zijn postzegels koopt. Als voorbeeld nemen we het volgende schema (in miljarden guldens per jaar), dat ruwweg in overeenstemming is met de Nederlandse verhoudingen omstreeks 1955:[1]

1. De getallen van dit schema zijn zodanig aangepast, dat de berekeningen die erop gebaseerd zijn redelijk eenvoudig blijven.

Verkopers (leveranciers)	Kopers (afnemers)				
	Landbouw	Industrie	Diensten	Consument	Totaal
Landbouw	1	2,25	0,2	1,55	5
Industrie	2	6	1	16	25
Diensten	0,2	3	1,8	15	20

Dit is een zgn. aan- en afvoertabel (Engels: input-output table), die specificeert tot welke bedragen de diverse sectoren aan goederen en diensten van andere sectoren betrekken (laten aanvoeren) en tot welke bedragen zij aan de verschillende sectoren leveren (afvoeren). Zo zien we, dat de totale jaarlijkse landbouwproductie 5 miljard bedroeg, dat daarvan 1 miljard, dus 20%, van de ene boer naar de andere ging; dat bijna de helft (2,25 miljard) geleverd werd aan de (voedselverwerkende) industrie; en dat bijna al het overige rechtstreeks naar de consument ging. Evenzo zien we, dat de sector industrie voor 2 miljard aan de landbouwsector leverde, dus iets minder dan de daaraan tegengestelde stroom van leveranties van de landbouw aan de industrie; dat de rechtstreekse leveringen van de dienstensector aan de consument (vervoersdiensten, handelsdiensten en ook de diensten van de radiodirigent) 15 miljard bedroegen; enz. De problemen die rijzen bij het vaststellen van de transactiebedragen moet men niet onderschatten. De belangrijkste bron van informatie wordt gevormd door productiestatistieken e.d. In beginsel zijn er voor elke transactie twee gegevens beschikbaar: het bedrag waarvoor de leverende sector beweert geleverd te hebben en het bedrag waarvoor de kopende sector beweert gekocht te hebben. Zijn beide gegevens beschikbaar èn met elkaar in overeenstemming, dan is dit de mooiste situatie die men zich kan wensen. Vaak echter zijn ze niet met elkaar in overeenstemming; dan dient men ze op redelijke wijze aan elkaar aan te passen. Soms is maar één van de twee gegevens beschikbaar; dan valt er weinig aan te passen. En soms geen van beide; of er zijn niet meer dan fragmentarische gegevens beschikbaar. Dan moeten intuïtie en ervaring het werk doen om nog tot een redelijke raming te komen. Het is duidelijk dat de resultaten met afwijkingen behept plegen te zijn.

Uit het transactieschema blijkt voorts, dat de consumenten in totaal kopen voor een waarde van 1,55 + 16 + 15 = 32,55 miljard gulden. Blijft dus de vraag waar de consumenten het

geld vandaan halen om landbouwproducten, industriegoederen en diensten ter waarde van 32,55 miljard gulden te kunnen kopen. Dit gebeurt, natuurlijk, met behulp van de lonen die de boerenknecht in de landbouw, de voorman in de industrie en de ambtenaar in de dienstensector verdient. En verder met behulp van de winst van de zelfstandige ambachtsman, de renten en dividenden van de rentenier en de kapitalist, enz. Kortom, met behulp van het inkomen dat in de desbetreffende sectoren wordt verdiend. Hoe berekenen we dit inkomen per sector? Neem de landbouw; in deze sector is de totale productie 5 miljard, waarvan 3,2 miljard direct kan worden verklaard. Voor die totale productie is het immers noodzakelijk, dat de landbouwbedrijven voor 1 miljard van andere landbouwbedrijven kopen, voor 2 miljard van industriële bedrijven en voor 0,2 miljard van de dienstensector; kortom, het gaat om de getallen van de eerste kolom, tezamen 3,2 miljard. Resteert 5 — 3,2 = 1,8 miljard. Dit is het bedrag dat de landbouwsector ontvangt minus het bedrag dat door die sector wordt afgestaan aan de drie sectoren, dus het is het bedrag dat beschikbaar blijft voor hen die hun krachten aan die sector geven: boerenknecht, boer, pachtheer. Het is het inkomen van de landbouwsector. Het schema ziet er dan als volgt uit:

Verkopers (leveranciers)	Kopers (afnemers)				
	Landbouw	Industrie	Diensten	Consument	Totaal
Landbouw	1	2,25	0,2	1,55	5
Industrie	2	6	1	16	25
Diensten	0,2	3	1,8	15	20
Sectorinkomen	1,8	13,75	17		32,55
Totaal	5	25	20	32,55	

De tweede kolom laat zien, dat de industrie van de landbouw, de industrie en de dienstensector betrekt voor 2,25 resp. 6 resp. 3 miljard, dus 11,25 miljard in totaal; door dit af te trekken van de totale industriële productie ad 25 miljard (het tweede getal van de laatste kolom) krijgen we dan het inkomen dat door arbeider, directeur en aandeelhouder in de industrie verdiend wordt, nl. 13,75 miljard. Op overeenkomstige wijze is het sectorinkomen van de dienstensector 17 miljard gulden.

Uit dit schema blijkt, dat het totale inkomen gelijk is aan

1,8 + 13,75 + 17 = 32,55 miljard, dus precies voldoende om te betalen wat er geconsumeerd wordt. We zien ook, dat het inkomen van de dienstensector (17 miljard) aanzienlijk groter is dan het totale bedrag waarvoor de bedrijven van die sector van andere bedrijven kopen (3 miljard). Dit is niet anders dan men op grond van directe ervaring zou verwachten. Een dienstverleningsbedrijf is immers 'arbeidsintensief', d.w.z. de lonen vormen een bijzonder belangrijke component van de kostenrekening.

2. AAN- EN AFVOERTABELLEN IN DE PRACTIJK

Om misverstand te vermijden dienen we uitdrukkelijk te vermelden, dat we in het bovenstaande een zeer vereenvoudigde voorstelling van zaken hebben gegeven. Zo is bijvoorbeeld de Nederlandse volkshuishouding bepaald niet wat men noemt een 'gesloten economie', d.w.z. een stelsel zonder verkeer met het buitenland. Er is dus ook sprake van import en export, en wel tot bedragen die zeker niet verwaarloosbaar gering zijn in vergelijking met wat er in het binnenland omgaat. Een tweede punt waarmee geen rekening werd gehouden vormen de besparingen; al het sectorinkomen ad 32,55 miljard werd volgens het schema ook besteed. Ook de overheid en de belastingen verdienen meer aandacht dan zij kregen. Met al deze verfijningen kan echter op in beginsel betrekkelijk eenvoudige wijze worden rekening gehouden. Wij zullen dit niet doen, omdat het dan snel erg technisch wordt.

Wél willen wij, alvorens op de theorie van de aan- en afvoeranalyse in te gaan, iets laten zien van de Nederlandse aan- en afvoertabellen, zoals die geconstrueerd zijn door het Centraal Bureau voor de Statistiek (C.B.S.) in Den Haag. Dit maakt het inzicht nog wat concreter en bovendien is er alle reden om aan juist de Nederlandse tabellen enige aandacht te besteden, want Nederland is het enige land dat die tabellen, al zijn ze niet zeer gedetailleerd,[1] *elk* jaar (sinds 1948) publiceert. Er worden 35 sectoren onderscheiden:

1. Wanneer de tabellen zeer gedetailleerd worden gemaakt, komt men in een klein land als Nederland al gauw in een situatie die strijdig is met de geheimhoudingsplicht. Men houdt bijv. alleen Philips als sector over (of evt. Philips plus een paar kleine bedrijven die door Philips overschaduwd worden, zodat ook dan het wel en wee van Philips uit de cijfers blijkt).

1. Landbouw, bosbouw en visserij
2. Steenkolenmijnen
3. Olie- en zoutwinning, veenderijen e.d.
4. Voedingsmiddelenindustrie (veehouderijproducten)
5. Voedingsmiddelenindustrie (overige producten)
6. Voortbrenging van dranken en tabaksproducten
7. Textielnijverheid
8. Schoeisel- en kledingindustrie
9. Hout- en meubelindustrie
10. Papiernijverheid
11. Drukkerijen en uitgeversbedrijven
12. Leder- en rubbernijverheid
13. Chemische nijverheid; raffinaderijen van petroleum
14. Vervaardiging van aardewerk, glas, kalk en stenen
15. Metallurgische industrie
16. Vervaardiging van metaalproducten en machinebouw
17. Electrotechnische industrie
18. Transportmiddelenindustrie
19. Overige metaalverwerkende bedrijven en diamantindustrie
20. Bouwnijverheid
21. Electriciteits-, gas- en waterleidingbedrijven
22. Groothandel
23. Kleinhandel
24. Banken en giro-instellingen
25. Verzekeringswezen
26. Woningbezit
27. Zeescheepvaart en luchtvaart
28. Overige vervoersbedrijven
29. Communicatiebedrijven
30. Medische en gezondheidsdiensten
31. Vrije beroepen en niet elders genoemde bedrijven
32. Vermakelijkheidsinstellingen
33. Hotels, café's, restaurants e.d.
34. Overige persoonlijke diensten
35. Goederen en diensten niet naar herkomst (resp. bestemming) verdeeld.

Zoals U ziet, een indrukwekkende lijst. De industrie in engere zin neemt rond de helft van de sectoren voor zijn rekening. Er zijn talrijke sectoren waarover het een en ander zou kunnen

worden gezegd (zoals No. 22 en 23, groot- en kleinhandel, waarvan de diensten gewaardeerd worden tegen de desbetreffende winstmarges), maar ook dat zou te technisch worden. We maken één uitzondering, nl. voor No. 35 (goederen en diensten niet naar herkomst resp. bestemming verdeeld). Wanneer zo'n tabel bij stukjes en beetjes wordt opgebouwd en tenslotte 'vol' is, moeten de geproduceerde en de afgezette hoeveelheden (afgezien van eventuele voorraadveranderingen) met elkaar in overeenstemming zijn. Dit nu is een onbereikbaar ideaal. We zagen al eerder dat de diverse getallen met fouten behept zijn, vandaar dat de totalen in feite niet kloppen. Om dit aan te geven wordt een laatste sector toegevoegd (hier No. 35), die niets anders is dan een restpost. Deze post is relatief onbelangrijk maar boekhoudkundig niet te vermijden.

Met 35 sectoren wordt de tabel natuurlijk veel groter dan met 3. De tabel waarmee wij begonnen op blz. 79 had aan $3 \times 3 = 9$ getallen voldoende om alle onderlinge leveringen te specificeren; hier zijn het er $35 \times 35 = 1225$. Om het een en ander nog iets concreter te maken zonder in een oerwoud van cijfers te verzeilen zullen wij een klein stukje van de tabel laten zien. Het betreft de onderlinge leveringen van de sectoren 6 t/m 10 (hieronder verkort omschreven) in miljoenen guldens per jaar voor 1957:

	6	7	8	9	10
6. Dranken, tabaksproducten	12	0	0	0	0
7. Textiel	0	785	332	21	4
8. Schoeisel, kleding	0	2	53	0	0
9. Hout, meubels	18	5	4	78	4
10. Papier	16	17	8	5	249

In totaal dus $5 \times 5 = 25$ onderlinge leveringen, waarvan 8 nullen. Die nullen zijn geen 'echte' nullen. De getallen zijn nl. afgerond tot op 1 miljoen nauwkeurig, zodat bijv. de 12 linksboven in beginsel betekent: tussen $11^{1}/_{2}$ en $12^{1}/_{2}$ miljoen gulden per jaar. (In beginsel, want we weten dat de getallen ook zonder afronding niet volmaakt nauwkeurig zijn.) Een nul betekent dus hier: minder dan $^{1}/_{2}$ miljoen per jaar. Het spreekt ook vanzelf, dat er wel enige leveringen zijn van de sector van dranken en tabaksproducten naar de vier andere, bijv. de leveranties aan de cantines van de spinnerijen (die vallen onder sector 7).

Wat opvalt is dat de leveringen van een sector aan dezelfde sector aan de hoge kant zijn: 12 miljoen voor sector 6, 785 miljoen voor sector 7, 53 miljoen voor sector 8, enz. (Dit zijn de zgn. diagonale elementen van de tabel.) Dat is niet onbegrijpelijk. In de textielnijverheid bijv. komt de katoen als import binnen. Volgens dezelfde tabel voor 1957 (maar niet in ons verkleinde overzicht vermeld) gaat van de totale Nederlandse import bijna 1 miljard naar de textielindustrie. Deze katoen gaat naar de spinnerijen. Van de spinnerijen gaat de katoen naar de weverijen of naar de tricotagefabrieken. In beide gevallen is dit een levering *van* de textielnijverheid *aan* de textielnijverheid. Van de weverijen gaat de textiel naar de zgn. 'finishing' industrieën, waar de lappen katoen worden geverfd, gebleekt, bedrukt, geborduurd, enz. Ook dit is weer een levering van de textielnijverheid aan de textielnijverheid. Meer technisch uitgedrukt: leveringen *binnen de sector* vormen bijdragen tot de op de diagonaal vermelde waarden. Op deze wijze zijn de relatief grote bedragen op de diagonaal te verklaren. Overigens ziet men, dat de leveringen van de textielnijverheid aan de schoeisel- en kledingindustrie (332 miljoen) bepaald ook niet te verwaarlozen zijn.

Wanneer we verder zouden gaan met de opsplitsing, dus meer sectoren zouden onderscheiden die dan elk kleiner zijn, dan zouden de bedragen op de diagonaal snel afnemen. Stel bijv. dat we niet met 'textielnijverheid' zouden werken maar met spinnerijen, weverijen, 'finishing' fabrieken en tricotagefabrieken als afzonderlijke groepen. Dan zouden de leveranties van spinnerijen aan weverijen buiten de diagonaal komen. (Die van spinnerijen aan spinnerijen blijven op de diagonaal staan, maar dat heeft niet veel te betekenen.) Ook grote industriële concentraties hebben dit effect. Wanneer bijv. in een en dezelfde onderneming èn gesponnen èn geweven èn bedrukt wordt, dan hebben de katoenleveringen van de spinnerijafdeling aan de weverijafdeling en verder binnen hetzelfde bedrijf plaats. Dergelijke leveringen binnen bedrijven worden niet in de aan- en afvoertabel opgenomen. Vandaar dat de leveringen van een sector aan dezelfde sector geringer worden naarmate de bedrijfseenheden waaruit de sector bestaat groter worden.

3. DOEL EN VERONDERSTELLINGEN VAN DE AAN- EN AFVOERANALYSE

Aan- en afvoertabellen van enige omvang hebben het grote voordeel, dat zij met een vrij grote mate van detail inzicht geven in de vervlechtingen van de diverse sectoren van het bedrijfsleven. Dat is natuurlijk prettig, maar de voor de hand liggende vraag is dan of dat inzicht ons ook in staat stelt op belangrijke vragen een goed antwoord te geven – althans een beter antwoord dan mogelijk is zonder de tabellen, die per slot van rekening met de nodige offers aan geld en mankracht zijn vervaardigd. Laten we daartoe stellen, dat U geïnteresseerd bent in de vraag, hoe groot de totale productie van de textielnijverheid over drie jaren zal zijn. U kunt bijv. een fabriek in Tilburg of Twente hebben staan en zekere ideeën hebben over de ontwikkeling van Uw marktaandeel in de nabije toekomst; dan is het bepaald nuttig om iets over die totale productie te weten, omdat U dan aan de hand van het marktaandeel berekeningen voor Uw eigen bedrijf kunt maken. U kunt een vakbondsbestuurder op textielterrein zijn; de ontwikkeling van de totale productie zal U dan interesseren, omdat die rechtstreeks verband houdt met de werkgelegenheid in die sector. En U kunt ambtenaar in Den Haag zijn, in het bijzonder belast met het wel en wee van die sector.

De vraag luidt dus: stelt een aan- en afvoertabel, die de textielnijverheid als een van de sectoren bevat, ons in staat de totale productie van die sector drie jaren vooruit te bepalen? Het antwoord luidt: als we de tabel op de juiste wijze hanteren én bereid zijn een aantal veronderstellingen te maken, ja! Als de gemaakte veronderstellingen dan (bij benadering) juist zijn, zullen we inderdaad de totale productie in de textielsector over drie jaar (bij benadering) correct kunnen bepalen. Hoe moeten we de tabel hanteren? Welke veronderstellingen moeten we maken?

In dit type van onderzoek wordt steeds als gegeven aangenomen tot welke bedragen de consument besluit van de diverse sectoren te consumeren. Zoals eigenlijk al uit de constructie bleek, vestigen de aan- en afvoertabellen in de eerste plaats de aandacht op de leveringen van de diverse sectoren aan elkaar. Wat daarbuiten gebeurt, dus met name welke bedragen de consumenten besteden aan de producten van die sectoren, wordt

als gegeven aangenomen. In dit geval gaan we er dus van uit dat we ons tevoren een idee hebben gevormd (bijv. m.b.v. een uitvoerig marktanalytisch onderzoek) over de uitgaven van de consumenten aan de producten van de verschillende sectoren over drie jaren. Dit is, zo U wilt, al één veronderstelling.

Er is nog een belangrijke veronderstelling. Om deze te illustreren stappen wij af van de textielnijverheid en de C.B.S.-tabellen, want het werken met 35 sectoren maakt de zaak ingewikkelder dan voor ons doel noodzakelijk is. Wij keren terug naar de drie sectoren van de tabel op blz. 79, die we hier ter wille van de overzichtelijkheid reproduceren:

Verkopers (leveranciers)	Kopers (afnemers)				
	Landbouw	Industrie	Diensten	Consument	Totaal
Landbouw	1	2,25	0,2	1,55	5
Industrie	2	6	1	16	25
Diensten	0,2	3	1,8	15	20

Het met het textielvoorbeeld overeenkomende probleem luidt nu voor dit gecomprimeerde schema: gegeven is dat de consument bepaalde bedragen besteedt aan de producten van de drie sectoren, te bepalen zijn de bedragen waarvoor bijv. de landbouwsector of de industrie dan in totaal moet produceren.

Beschouw de landbouw. De totale productie van alle bedrijven in die sector bedraagt 5 miljard; en om die totale productie mogelijk te maken is het blijkens de tabel in het jaar in kwestie noodzakelijk geweest, dat die bedrijven van andere landbouwbedrijven voor 1 miljard betrokken, van industriële bedrijven voor 2 miljard en van de dienstensector voor 0,2 miljard. Hoe zal het zijn wanneer de totale productie van de landbouwsector niet 5 maar 6 miljard is, dus 20% hoger? Het ligt in de rede, dat de verwezenlijking van dit omvangrijker programma een grotere aanvoer van goederen en diensten van andere bedrijven vergt. De eenvoudigste relatie is die van evenredigheid. D.w.z., *stijgt de totale productie van de landbouwsector met 20%, dan dient de landbouwsector van ieder der sectoren 20% meer geleverd te krijgen.* Dus 1,2 miljard i.p.v. 1 miljard van de land-

bouw, 2,4 i.p.v. 2 miljard van de industrie en 0,24 i.p.v. 0,2 miljard van de dienstensector. Onrealistisch? Zonder twijfel. Niemand zal geloven, dat die 20% extra totale productie zal doorwerken in precies dezelfde percentages voor alle aan de landbouw leverende sectoren. Maar wie volmaakte precisie wil komt niet ver op het terrein van de economie. Wij moeten hopen op en streven naar redelijke benaderingen – meer niet.

Laten we nog even doorgaan met die 20% verhoging. Vòòr de verhoging hadden we 5 miljard totale landbouwproductie en 1 resp. 2 resp. 0,2 miljard voor de leveringen van de drie sectoren aan de landbouw. Delen we die drie getallen door de totale landbouwproductie, dan krijgen we 0,2 resp. 0,4 resp. 0,04. Dit zijn dus de leveringen van de drie sectoren aan de landbouw *per* miljard totale landbouwproductie. Na de verhoging ad 20% wordt de totale landbouwproductie 6 miljard; aannemende dat de leveringen ook met elk 20% stijgen, dan worden zij zoals wij zagen 1,2 resp. 2,4 resp. 0,24 miljard. Gaan we in de nieuwe situatie die leveringen delen door de totale landbouwproductie, dan krijgen we

$$\frac{1,2}{6} = 0,2 \quad \frac{2,4}{6} = 0,4 \quad \frac{0,24}{6} = 0,04,$$

dus precies dezelfde verhoudingen als tevoren! Kennelijk komt de veronderstelling van gelijke procentuele toenemingen neer op constante verhoudingen, preciezer: op een onveranderd blijven van de verhouding van de leveringen van elk der drie sectoren aan de landbouw tot de totale landbouwproductie. Dit is eenvoudig het gevolg van onze veronderstelling.

Laten we nu de twee andere sectoren nog even bezien. We hebben de leveringen aan de landbouw gedeeld door de totale landbouwproductie. Op overeenkomstige wijze gaan we de leveringen aan de industrie delen door de totale productie van de industrie en de leveringen aan de dienstensector door de totale productie van die sector. Zo krijgen we bijv. 2,25/25 = 0,09 voor de verhouding van de leveranties van de landbouw aan de industrie tot de totale industriële productie. Nemen we weer aan dat die verhouding constant is, dan betekent dit dat wat de landbouw aan de industrie levert *steeds* 9% bedraagt van wat de industrie in totaal produceert. Zo komen we tot het volgende overzicht:

Verkopers (leveranciers)		Kopers (afnemers)	
	Landbouw	Industrie	Diensten
Landbouw	0,2	0,09	0,01
Industrie	0,4	0,24	0,05
Diensten	0,04	0,12	0,09

De getallen van deze staat zijn dus als volgt afgeleid: we bepalen voor *elke* sector de behoeften aan leveringen door *alle* sectoren per eenheid *totale productie* van de sector; dus we delen de getallen van de eerste kolom van de tabel op blz. 86 door het eerste element van de totaalkolom (hetgeen leidt tot de eerste kolom van deze staat) en we delen de getallen van de tweede kolom van de tabel op blz. 86 door het tweede element van de totaal-kolom (het resultaat is de tweede kolom van deze staat) en analoog voor de derde kolom. De negen coëfficiënten van deze staat staan wel bekend als 'technische coëfficiënten'. Waarmee dan wordt aangegeven, dat zij de behoefte aan leveranties preciseren op grond van het feit, dat het nu eenmaal technisch noodzakelijk is om grondstoffen en diensten aan een sector te leveren als die sector voor een gegeven bedrag dient te produceren. De term suggereert, dat die verhoudingen op technische gronden vastliggen, d.w.z. dat het een technische wet is dat voor 10% meer afvoer steeds 10% meer aanvoer van alle leveranciers noodzakelijk is. Dat is natuurlijk overdreven. De veronderstelling van vaste coëfficiënten is een eerste benadering van de werkelijkheid, meer niet.

4. HOE WERKT DE AAN- EN AFVOERANALYSE?

Het gaat nog steeds hierom: de totale productie per sector te bepalen, gegeven de bedragen die de consument per sector besteedt. Het middel om tot dit doel te komen: een aan- en afvoertabel van een zeker jaar plus de veronderstelling dat de technische coëfficiënten van dat jaar dezelfde zijn als die van het jaar waarvoor wij de productie per sector willen bepalen. Of ook: de veronderstelling dat voor elke sector de behoefte aan leveringen door alle sectoren per eenheid totale productie in de loop van de tijd constant blijft.

Stel dat om wat voor reden dan ook de consumptie van land-

bouwproducten zou stijgen met 0,5 miljard, dus van 1,55 miljard tot 2,05 miljard. We bepalen ons tot een verandering van de consumptie van deze ene sector; we laten dus de consumptie van industriële producten en die van de dienstensector op het oude niveau (16 resp. 15 miljard), want we willen het probleem zo eenvoudig mogelijk stellen. Ook de oplossing zullen we in eerste instantie op heel eenvoudige wijze afleiden, omdat we juist daardoor op heldere wijze zien hoe de samenhang van de verschillende sectoren van het bedrijfsleven doorwerkt.

Zou het voldoende zijn dat de landbouwsector 0,5 miljard extra produceert om te voldoen aan de extra consumptieve vraag van 0,5 miljard? Neen! Want de technische coëfficiënten laten zien, dat wil die extra productie tot stand komen, allereerst noodzakelijk is dat de landbouwbedrijven van andere landbouwbedrijven voor een bedrag van 0,2 × 0,5 = 0,1 miljard betrekken (0,2 is het getal linksboven in de tabel van technische coëfficiënten op blz. 88). Daarmee stijgt de extra productie nodig voor die extra consumptieve vraag tot 0,5 + 0,1 = 0,6. En daarmee zijn we er nog niet. Want zelfs al zou maar 0,5 miljard extra productie nodig zijn, dan is daarvoor noodzakelijk dat de landbouwbedrijven extra producten van de industrie en van de dienstensector geleverd krijgen. Aan de hand van de eerste kolom van de tabel van technische coëfficiënten zien we dat de extra leveranties van de industrie aan de landbouw 0,4 × 0,5 = 0,2 miljard bedragen en voor de dienstensector aan de landbouw 0,04 × 0,5 = 0,02 miljard. En dus, terwijl we eerst wellicht dachten dat het voldoende zou zijn de productie van de landbouwsector te verhogen met 0,5 miljard, zien we nu dat in feite aanmerkelijk meer moet worden geproduceerd: 0,6 miljard door de landbouwsector en 0,2 resp. 0,02 miljard door de industrie resp. de dienstensector. En dat terwijl de consument in het geheel geen beroep heeft gedaan op de laatste twee sectoren!

Maar we zijn er nog niet. Die extra behoeften van 0,1 miljard te leveren door de landbouw, van 0,2 miljard te leveren door de industrie en van 0,02 miljard te leveren door de dienstensector zijn alle berekend onder de veronderstelling dat het alleen maar gaat om een verhoogde totale landbouwproductie van 0,5 miljard. We weten nu wel beter! Om te beginnen hebben we nog eens 0,1 miljard productie extra van de landbouw; en die vereist op precies dezelfde manier:

0,2 × 0,1 = 0,02 miljard van de landbouw;
0,4 × 0,1 = 0,04 miljard van de industrie;
0,04 × 0,1 = 0,004 miljard van de dienstensector.

We hebben ook nog die 0,2 miljard extra te produceren door de industrie. Aan de hand van de tweede kolom van de tabel van technische coëfficiënten zien we, dat die vereist:

0,09 × 0,2 = 0,018 miljard van de landbouw;
0,24 × 0,2 = 0,048 miljard van de industrie;
0,12 × 0,2 = 0,024 miljard van de dienstensector.

En zo hebben we ook nog de extra 0,02 miljard te leveren door de dienstensector. De derde kolom van de tabel van technische coëfficiënten laat zien, dat dit leidt tot de volgende extra leveringen:

0,01 × 0,02 = 0,0002 miljard van de landbouw;
0,05 × 0,02 = 0,0010 miljard van de industrie;
0,09 × 0,02 = 0,0018 miljard van de dienstensector.

Zo hebben we dus voor elk van de drie sectoren drie bedragen waarvoor in deze ronde extra dient te worden geproduceerd. Het totaal van deze ronde is:

0,02 + 0,018 + 0,0002 = 0,0382 miljard van de landbouw;
0,04 + 0,048 + 0,0010 = 0,0890 miljard van de industrie;
0,004 + 0,024 + 0,0018 = 0,0298 miljard van de dienstensector.

De totale extra landbouwproductie is dus nu gestegen van 0,6 tot 0,6382 miljard; voor de industrie van 0,2 tot 0,2890 miljard; en voor de dienstensector van 0,02 tot 0,0498 miljard.

Maar daarmee zijn we er nog steeds niet, want de extra productie van deze ronde vergt op zijn beurt extra leveringen, en zo gaat het door! Ingewikkeld? Nogal. Komt er een eind aan al die leveringen? Ja en neen. *Neen* in de zin dat die keten van successieve leveringen, zoals wij die hier voorstellen, in beginsel 'tot in het oneindige' doorgaat. *Ja* in de zin dat het uiteindelijk toch maar om beperkte bedragen gaat. Een eerste indicatie in dit opzicht hebben we door de hierboven verkregen uit-

komsten. In eerste ronde ging het om een extra productie van de landbouw ten bedrage van 0,5 miljard. In tweede ronde ging het om 0,1 miljard van de landbouw, 0,2 miljard van de industrie en 0,02 miljard van de dienstensector, d.w.z. om 0,32 miljard in totaal; dus bijna 40% minder dan het bedrag van de eerste ronde. En in derde ronde gaat het, naar men gemakkelijk kan narekenen, om een bedrag dat ongeveer de helft minder is dan dat van de tweede ronde. Zo gaat dat verder. Men kan deze situatie het gemakkelijkst doorzien door hem te verplaatsen naar een wat meer alledaagse sfeer. Stel dat Uw zoon (10 jaar) U zijn nijpend geldgebrek komt uitleggen; U geeft hem een gulden. Een week later komt hij terug met hetzelfde probleem. Hij en U weten van geven en nemen en het resultaat is dat hij met een halve gulden heen gaat. En dat herhaalt zich week na week, elke keer krijgt hij de helft van wat hij de vorige week kreeg, tot in lengte van dagen. (Zo lang zult U niet leven en hij ook niet en het betalen van delen van centen stuit op practische bezwaren; maar U kunt het zich in abstracto voorstellen.) Dat betekent dat de reeks van betalingen zich tot in het oneindige uitstrekt. Maar het totale *bedrag* waar het om gaat is beperkt. Het is

$$1 + {}^{1}/_{2} + {}^{1}/_{4} + {}^{1}/_{8} + {}^{1}/_{16} + \ldots,$$

dus *nu* is er *in totaal* f. 1 uitbetaald, een week later f. 1,50, twee weken later f. 1,75, enz. Zo nadert U tot 2 gulden; U komt er willekeurig dicht bij naarmate de tijd verder voortschrijdt. Die 2 gulden is wat in de wiskunde een 'limiet' wordt genoemd.[1]

Zo gaat het ons ook hier om de limiet; d.w.z. om het bedrag waarvoor elk van de drie sectoren extra moet produceren teneinde te voldoen aan die extra consumptieve vraag naar landbouwproducten van 0,5 miljard, daarbij rekening houdend met de noodzaak van *al* die successieve onderlinge leveringen. Moeten de berekeningen zo moeizaam zijn als die van hierboven? Bepaald niet, het kan veel vlotter.

Laten we *L* schrijven voor de totale landbouwproductie (in miljarden guldens per jaar), *I* voor de totale industriële productie en *D* voor de totale productie van de dienstensector. We beschouwen nu eerst de aan- en afvoertabel op blz. 86. De eerste rij laat zien, dat *L* (daar gelijk aan 5) de som is van

1. U bereikt die 2 gulden nooit helemaal! Want

$$1 + {}^{1}/_{2} + {}^{1}/_{4} + {}^{1}/_{8} + {}^{1}/_{16} = 1{}^{15}/_{16},$$

een week later wordt het totaal $1{}^{31}/_{32}$, nog een week later $1{}^{63}/_{64}$, enz.

vier bedragen, nl. de leveranties van de landbouw aan de drie sectoren plus wat de consument van de landbouw koopt. Dit laatste bedrag is gegeven, nl. 2,05 miljard, want het gaat juist om de verhoging van het consumptieve verbruik van landbouwproducten van 1,55 tot die 2,05 miljard. Voor de eerste drie componenten van L, dus de leveranties van de landbouw aan de drie sectoren, gaan we naar de tabel van de technische coëfficiënten op blz. 88. De eerste rij (van achter naar voren bezien) specificeert, dat een totale productie van de dienstensector ten bedrage van D een levering ten bedrage van $0{,}01D$ van de landbouw aan de dienstensector vereist; dat als I de totale productie van de industrie is, de landbouw aan de industrie voor $0{,}09I$ moet leveren; en tenslotte, dat de landbouwbedrijven aan elkaar voor $0{,}2L$ moeten leveren. Zo komen we tot de volgende vergelijking:

$$L = 0{,}2L + 0{,}09I + 0{,}01D + 2{,}05.$$

We trekken $0{,}2L + 0{,}09I + 0{,}01D$ van linker- en rechterlid af en krijgen zo:

(1) $0{,}8L - 0{,}09I - 0{,}01D = 2{,}05.$

Vervolgens gaan we naar de tweede rij van de aan- en afvoertabel op blz. 86, die op de industrie betrekking heeft. De totale productie (I) bestaat ook nu uit vier componenten. De consumptie bedraagt 16 miljard; we spraken immers af om deze ongewijzigd op het oude niveau te laten. De drie andere componenten hebben betrekking op hetgeen de industrie aan de drie sectoren moet leveren, gegeven het feit dat deze sectoren een totale productie van L resp. I resp. D hebben te verzorgen. De tabel van technische coëfficiënten (tweede rij) op blz. 88 laat dan zien, dat aan de landbouw voor $0{,}4L$ geleverd moet worden, aan de industrie voor $0{,}24I$ en aan de dienstensector voor $0{,}05D$. En zo hebben we dan:

$$I = 0{,}4L + 0{,}24I + 0{,}05D + 16$$

of na aftrek van $0{,}4L + 0{,}24I + 0{,}05D$ links en rechts:

(2) $-0{,}4L + 0{,}76I - 0{,}05D = 16.$

Analoog voor de dienstensector:

$$D = 0{,}04L + 0{,}12I + 0{,}09D + 15$$

of ook:

(3) $-0{,}04L - 0{,}12I + 0{,}91D = 15.$

Bekijk nu de vergelijkingen (1), (2) en (3). Dit zijn drie lineaire vergelijkingen in de drie onbekenden L, I en D. En oplosbaar! De oplossing luidt:

$L = 5{,}6659$
$I = 25{,}3555$
$D = 20{,}0762$

zoals men gemakkelijk verifieert door de oplossingen in de vergelijkingen in te vullen.[1] Wanneer we in herinnering roepen, dat de oorspronkelijke situatie gekenmerkt was door totale sectorproducties van resp. 5, 25 en 20 miljard (zie de tabel op blz. 86), dan kunnen we concluderen: wat uiteindelijk *extra* geproduceerd moet worden, gegeven het feit dat de consumenten voor een half miljard extra bij de landbouw kopen, is:

665,9 miljoen door de landbouw;
355,5 miljoen door de industrie;
76,2 miljoen door de dienstensector.

Daarmee is ons vraagstuk opgelost. Toen wij op de stap-voor-stap manier (de 'iteratieve' manier) te werk gingen, concludeerden we aan het einde van de derde ronde, dat de extra bedragen waren:

638,2 miljoen door de landbouw;
289,0 miljoen door de industrie;
49,8 miljoen door de dienstensector.

We waren dus wel een eind op weg, maar nog lang niet waar we moesten zijn.

5. NOG ENKELE DETAILS

De hier geschetste procedure van het oplossen van 3 vergelijkingen met 3 onbekenden is duidelijk veel eenvoudiger dan de oorspronkelijke iteratieve, waarbij we ronde na ronde de consequenties van de samenhang der sectoren moesten doorrekenen. Maar het kan nog veel vlotter. Dat is ook van groot belang, omdat het puur toeval zou zijn wanneer men nu net

1. Vul de gevonden waarden van L, I en D in het linkerlid van (1) in: $0{,}8 \times 5{,}6659 = 4{,}5327$; $0{,}09 \times 25{,}3555 = 2{,}2820$; $0{,}01 \times 20{,}0762 = 0{,}2008$. Het linkerlid wordt dan $4{,}5327 - 2{,}2820 - 0{,}2008 = 2{,}0499$, dus afgezien van een kleine afrondingsfout gelijk aan 2,05, het rechterlid van (1). Voor (2) en (3) gaat het analoog.

geïnteresseerd zou zijn in een verhoging van de consumptieve uitgaven aan landbouwproducten met een half miljard, niet meer en niet minder, en in ongewijzigde consumptieve uitgaven aan de producten van de twee andere sectoren. Natuurlijk, wanneer men een willekeurige andere specificatie van de consumptieve bestedingen zou maken gaat het net zo; men moet dan een ander stelsel van drie lineaire vergelijkingen in drie onbekenden oplossen. Dit kan men eventueel keer op keer herhalen. Op zichzelf is dit geen ramp wanneer er maar 3 sectoren zijn, maar het wordt onprettig wanneer we met 35 sectoren te doen hebben, zoals het geval is met de C.B.S.-tabellen. Zelfs een electronische rekenmachine van de middenklasse heeft daar nog een flinke kluif aan. Het is daarom van groot belang, dat er een methode bestaat die vrijwel onmiddellijk de oplossing geeft voor *alle denkbare bedragen* van de consumptieve uitgaven aan de verschillende sectoren.

Om te laten zien hoe die methode werkt gaan we terug naar de vergelijkingen (1), (2) en (3). Dan valt het op, dat de consumptieve uitgaven alle rechts van het gelijkteken voorkomen: 2,05 in de landbouwvergelijking (1), 16 in de industrievergelijking (2) en 15 in de dienstenvergelijking (3). Links staan steeds de onbekenden L, I en D, elk vermenigvuldigd met zekere coëfficiënten. Die coëfficiënten nu zijn alle afgeleid uit de tabel van technische coëfficiënten. Maar die tabel verandert niet – dat is onze veronderstelling! Kortom, hoe we de consumptieve bestedingen ook specificeren, altijd blijven de coëfficiënten links dezelfde maar die rechts veranderen naar de aard van de specificatie. Laten we dan C_L schrijven voor de consumptieve uitgaven aan landbouwproducten, C_I voor die aan industriële producten en C_D voor die aan de producten van de dienstensector. De vergelijkingen (1), (2) en (3) krijgen nu de volgende vorm:

$$0{,}8L - 0{,}09I - 0{,}01D = C_L$$
$$-0{,}4L + 0{,}76I - 0{,}05D = C_I$$
$$-0{,}04L - 0{,}12I + 0{,}91D = C_D.$$

Ook dit stelsel kan men oplossen! De oplossing is:

$$L = 1{,}3319C_L + 0{,}1614C_I + 0{,}0235C_D$$
$$I = 0{,}7110C_L + 1{,}4135C_I + 0{,}0855C_D$$
$$D = 0{,}1523C_L + 0{,}1935C_I + 1{,}1112C_D.$$

Met behulp van deze vergelijkingen kunnen we nu, *gegeven* C_L, C_I en C_D, onmiddellijk vaststellen hoeveel er in totaal in ieder der sectoren moet worden geproduceerd. Het probleem is

tot een invuloefening teruggebracht. Stellen we bijvoorbeeld $C_L = 1{,}55$ en $C_I = 16$ en $C_D = 15$, dan vinden we de oplossing:

$$L = 1{,}3319 \times 1{,}55 + 0{,}1614 \times 16 + 0{,}0235 \times 15 = 5{,}00$$
$$I = 0{,}7110 \times 1{,}55 + 1{,}4135 \times 16 + 0{,}0855 \times 15 = 25{,}00$$
$$D = 0{,}1523 \times 1{,}55 + 0{,}1935 \times 16 + 1{,}1112 \times 15 = 20{,}00$$

Dit zal ons niet verbazen – wij wisten dit al van het transactieschema dat ons uitgangspunt vormde, zie blz. 86. We kunnen ook C_L, C_I en C_D andere waarden geven: het blijft een invuloefening.

Het zal intussen duidelijk zijn, dat in dit zeer sterk gecomprimeerde model slechts bepaald kan worden hoeveel in totaal in de industriële sector moet worden geproduceerd; dus niet bijvoorbeeld afzonderlijk de vereiste productie in de textielnijverheid en in de papierbranche. Maar de C.B.S.-tabel, die een veel fijnere indeling heeft, kan dat wél. Keren wij terug naar de vraag hoeveel er in de textielnijverheid over 3 jaar moet worden geproduceerd, en naar de C.B.S.-tabel. Wij moeten dan als *gegeven* aannemen de totale vraag over 3 jaar in ieder der 35 sectoren: zeg $C_1, C_2, \ldots, C_{35}$. Volgens de lijst van sectoren vermeld op blz. 82 staat C_7 voor de totale vraag (over 3 jaar) van de consumenten naar de producten van de textielnijverheid, C_8 voor de totale vraag (over 3 jaar) van de consumenten naar schoeisel en kleding, enz. Naast de veronderstelling dat het consumptiepatroon $C_1, C_2, \ldots, C_{35}$ bekend is, worden de technische coëfficiënten bekend verondersteld. Deze coëfficiënten worden weer bepaald met behulp van de aan- en afvoertabel: elk getal van de eerste kolom wordt gedeeld door de bijbehorende totale productie van de kopende sector, dus het eerste element van de totaal-kolom, idem voor de tweede, enz., tot de 35ste toe. Op deze wijze worden $35 \times 35 = 1225$ technische coëfficiënten berekend. Wij reproduceren hieronder de 25 coëfficiënten (met 100 vermenigvuldigd), die betrekking hebben op de sectoren 6 t/m 10; deze sectoren hebben we op blz. 83 al in detail beschouwd.

	6	7	8	9	10
6. Dranken, tabaksproducten	0,90	0	0	0	0
7. Textiel	0	27,30	19,73	2,28	0,37
8. Schoeisel, kleding	0	0,07	3,15	0	0
9. Hout, meubels	1,35	0,17	0,24	8,46	0,37
10. Papier	1,20	0,59	0,48	0,54	23,08

De getallen van de laatste kolom (onder 10 — Papier) zijn dus verkregen door de leveringen aan de papiersector (0, 4, 0, 4, 249 miljoen volgens de aan- en afvoertabel van blz. 83) te delen door de waarde van de totale productie van die sector[1] en door vervolgens te vermenigvuldigen met 100. De eigenlijke technische coëfficiënten van die sector zijn dus: 0; 0,0037; 0; 0,0037; 0,2308. En analoog voor de andere kolommen. Door de vermenigvuldiging met 100 zijn de technische coëfficiënten in feite als procenten uitgedrukt.

Uit de getallen van de eerste kolom (onder 6) blijkt dan, dat voor de productie van 100 gulden aan dranken en tabaksproducten nodig is 0,90% van die 100 gulden oftewel 90 cent aan andere dranken en tabaksproducten, 1,35% of f 1,35 aan hout en meubels (kratten voor de bierflessen!) en 1,20% of f. 1,20 aan papier (cigarettendoosjes, cigarettenpapier). Maar geen textiel en geen schoeisel en geen kleding. Het zijn deze getallen welke bekend worden verondersteld; zij worden *constant* verondersteld in de loop van de tijd. Het is nuttig deze technische coëfficiënten te vergelijken met de bedragen in het transactieschema zelf, dat werd vermeld op blz. 83. Als er geen (directe) transacties zijn tussen twee sectoren, zijn de technische coëfficiënten automatisch nul. Dit is bijv. het geval voor leveringen van de sector textiel aan de sector dranken en tabaksproducten, zoals wordt aangegeven door de bovenste nul (het tweede getal) in de eerste kolom van dat transactieschema.

Gewapend met al deze gegevens – de technische coëfficiënten én het consumptiepatroon $C_1, C_2, \ldots, C_{35}$ over drie jaar – kunnen wij de vraag beantwoorden hoeveel er over drie jaar in de textielbranche zal moeten worden geproduceerd. Daartoe dienen we echter een stelsel van 35 lineaire vergelijkingen in 35 onbekenden op te lossen, zoals wij op blz. 92 een vergelijkingenstelsel van drie vergelijkingen in drie onbekenden moesten oplossen. Gelukkig heeft het C.B.S. deze berekeningen al voor ons uitgevoerd. Als wij T schrijven voor de totale waarde van de productie in de sector textiel nodig om het gegeven consumptiepatroon te kunnen realiseren, dan valt uit één van de C.B.S.-tabellen af te lezen:

$$100T = 0{,}53C_1 + \ldots + 0{,}11C_6 + 137{,}67C_7 + 28{,}65C_8 + 3{,}56C_9 + 0{,}81C_{10} + \ldots + 0{,}32C_{35},$$

1. De totale productiewaarde van de papiersector in 1957 bedraagt 1079 miljoen gulden.

in woorden: laten de consumenten aan de producten van de eerste sector voor 100 gulden besteden, dan is daartoe nodig dat de textielsector voor 53 cent[1] produceert; ...; laten de consumenten aan de producten van de zesde sector (dranken en tabaksproducten) voor 100 gulden besteden, dan is daartoe nodig dat de textielsector voor 11 cent produceert; laten de consumenten aan de producten van de zevende sector, dus de textielsector *zelf*, 100 gulden besteden, dan is daartoe noodzakelijk dat die sector voor f. 137,67 produceert; enz. En wanneer het niet om 100 gulden gaat maar om een willekeurig veelvoud daarvan, dan wordt de daartoe noodzakelijke productie van de textielsector gevonden door bovengenoemde uitkomsten evenredig op te hogen. En wanneer de consumenten aan al die sectoren tegelijkertijd hun geld uitgeven, zoals in feite natuurlijk steeds het geval is, dan wordt de totale daartoe benodigde textielproductie gevonden door al die afzonderlijke componenten op te tellen.

Het verbaast U misschien dat C_6 in deze vergelijking voorkomt met een coëfficiënt 0,11. Dit betekent, zoals wij zojuist zagen, dat als men voor 100 gulden aan dranken en tabaksproducten wil consumeren, voor 11 cent aan textielgoederen geproduceerd moet worden. Dit lijkt een beetje merkwaardig, want de productie van dranken en tabak vereist in eerste aanleg helemaal geen productie van textiel: de technische coëfficiënt is nul. Maar *indirect* is er wel textiel nodig. Ook dit kan uit het schema van technische coëfficiënten direct worden afgeleid. Bijvoorbeeld, 100 gulden aan dranken en tabaksproducten vereist f. 1,35 aan hout en meubelproducten. Maar de f. 1,35 aan hout en meubelproducten, sector 9, bevat voor 2,28% of ongeveer 3 cent aan textiel! En analoog: 100 gulden productie in sector 6 vereist f. 1,20 van sector 10, doch f. 1,20 van sector 10 vereist 0,37% of bijna een halve cent van sector 7, de textielnijverheid. Ook hier weer een indirect verband. Al deze indirecte behoeften aan textiel samen (die ook best nóg ingewikkelder, nóg indirecter kunnen zijn) bedragen blijkbaar 11 cent.

Laten we nog iets algemener te werk gaan door aan te nemen dat men 100 gulden wenst te consumeren in *ieder* der sectoren 6 t/m 10. Dus $C_6 = C_7 = C_8 = C_9 = C_{10} = 100$ gulden. De

1. 53 cent, niet 53 gulden. Dat komt omdat in het linkerlid $100T$ staat i.p.v. T.

hiervoor benodigde textielproductie heeft volgens de zojuist besproken vergelijking voor $100T$ een waarde van

0,11 (voor C_6)
137,67 (voor C_7)
28,65 (voor C_8)
3,56 (voor C_9)
0,81 (voor C_{10})

dus in totaal f. 170,80. Van deze totale textielproductie komt slechts f. 100 de consumptie ten goede. De rest, dus f. 70,80, dient voor onderlinge leveringen. Het is interessant de componenten van dit bedrag aan onderlinge leveringen, dus

0,11	37,67	28,65	3,56	0,81
(Sector 6)	(Sector 7)	(Sector 8)	(Sector 9)	(Sector 10)

te vergelijken met de corresponderende tweede rij van de tabel van technische coëfficiënten, dus met

0	27,30	19,73	2,28	0,37.

Deze laatste getallen geven aan, dat als men voor een waarde van f. 100 produceert in elk van de sectoren 6 t/m 10, daartoe niet alleen noodzakelijk is dat die sectoren elk voor f. 100 produceren, maar ook dat zij elkaar onderling leveren; en in het bijzonder, wat de textielnijverheid betreft, dat deze sector voor 0 + 27,30 + . . . + 0,37 gulden aan dit vijftal sectoren levert. Maar die onderlinge leveringen werken verder door, ook via sectoren buiten dit vijftal; en wat de textielnijverheid uiteindelijk aan de vijf sectoren heeft te leveren is daarom aanzienlijk meer, nl. 0,11 + 37,67 + . . . + 0,81 gulden, zoals de eerste groep van getallen te zien geeft.

6. KNELPUNTEN

Tot nu toe zijn we ervan uitgegaan, dat de lezer om wat voor reden dan ook bijzondere belangstelling heeft voor het wel en wee van de textielbranche over drie jaar. Dit is natuurlijk een beperkte probleemstelling. In de Verenigde Staten zijn aan- en afvoertabellen in de Tweede Wereldoorlog gebruikt voor een probleem van veel grotere allure: het uitzoeken en vermijden van knelpunten. De snelle omschakeling naar oorlogsindustrie-

en en het op volle capaciteit bezet zijn van vele soorten arbeid, machines en andere productiemiddelen kan nl. leiden tot ernstige knelpunten. Het probleem is die zoveel mogelijk te vermijden en daarbij dient men dóór te denken. Als er op een scheepswerf op zeker moment voldoende mankracht is, kàn men besluiten een schip te bouwen. Maar zo'n schip vereist ijzer, en ijzer vereist kolen, en kolenmijnen vereisen arbeid. Als er nu onvoldoende mijnwerkers zijn kan het aanbeveling verdienen een deel van het personeel van de werf naar de mijnen te sturen. Het grote voordeel van de aan- en afvoeranalyse is dat hij dit soort indirecte effecten op duidelijke wijze laat zien.

Wij zullen, bij wijze van slot, ook hiervan een eenvoudige illustratie geven. Daartoe keren we terug naar de simpele tabel met drie sectoren: landbouw, industrie, diensten. Voorgesteld wordt het volgende consumptieprogramma: $C_L = 6$, $C_I = 50$, $C_D = 60$, alle in miljarden guldens per jaar. Dus 3 à 4 keer zoveel als de consumptiebedragen van de tabel op blz. 86. Een dergelijke hoorn des overvloeds zou ieder ruimschoots voortreffelijk te eten geven; iedereen zou zich een auto kunnen veroorloven en artsen zouden volop aanwezig zijn. Het ermee corresponderende productieprogramma zou vereisen (een invuloefening aan de hand van de formules op blz. 94):

$L = 17{,}4714$
$I = 80{,}0710$
$D = 77{,}2608.$

Een pittig program, maar waarom niet? Er is een heel eenvoudig antwoord: het is geen haalbare kaart. Wij zagen immers (zie de tabel op blz. 80), dat per 5 miljard totale landbouwproductie er een sectorinkomen in de landbouw van 1,8 miljard is, dus 36%. Voor de industrie en de dienstensector zijn deze percentages nog hoger. Dat inkomen bestaat voor het grootste deel (ruwweg twee derden) uit arbeidsbeloning: loon voor de glasblazer, salaris voor de boekhouder, 15% voor de kellner. (Het restant is beloning van kapitaal en grond: rente, dividenden, pachten.) Willen we de totale productie zo drastisch verhogen als hier wordt voorgesteld, dan dient de totale hoeveelheid arbeid ook drastisch te worden verhoogd. Verhoging is denkbaar: groter aantal arbeiders, productiviteitsstijging. Maar de verandering waarom het hier gaat is zo groot, dat daar op korte termijn eenvoudig niet aan te denken is. Het programma $C_L = 1{,}6$; $C_I = 17$; $C_D = 16{,}5$ is een stuk reëler.

Het impliceert een stijging van het totale inkomen van 32,55 tot 35,1 miljard, dus ruim 7%. Een dergelijke stijging valt bijv. te bereiken met 5% productiviteitsverhoging en 2% stijging van het aantal arbeiders.

In dit voorbeeld hebben we een geval waar de arbeid het knelpunt vormt, dat de groei van de consumptie binnen zekere perken houdt. Wanneer er volledige werkgelegenheid is kan dit inderdaad juist zijn. Maar er kunnen ook andere knelpunten zijn: gebouwen, machines e.d., kortom kapitaalgoederen, buitenlands geld om importen te betalen, enz. Ook dit kan in beginsel met aan- en afvoeranalyse behandeld worden; we verwijzen naar de literatuur.

LITERATUUR

Er is een uitgebreide literatuur over aan- en afvoeranalyse. Het oorspronkelijke werk op dit terrein is dat van Leontief [1]. Een zeer goede behandeling wordt gegeven door Chenery and Clark [2]. Een uitvoerige Nederlandse tekst is geschreven door twee Belgische auteurs, Van Straelen en Virenque [3]. De Amsterdamse lector Venekamp heeft de aan- en afvoeranalyse (door hem inzet-afzet genoemd) toegepast op de 'economie' van Amsterdam [4]. De 'rest van Nederland' wordt dan, evenals andere landen, als 'buitenland' opgevat. In 1953 werd in Nederland een conferentie over de aan- en afvoeranalyse gehouden, waar bijna alle 'grote namen' aanwezig waren [5]. Voor de C.B.S.-tabellen, waarnaar in de tekst herhaaldelijk werd verwezen, zie [6].

[1] Leontief, W. W., *The Structure of the American Economy, 1919-1939.* Second Edition. Oxford University Press, New York. 1951. (De eerste druk verscheen in 1941.)

[2] Chenery, H. B., and P. G. Clark, *Interindustry Economics.* John Wiley & Sons, Inc., New York. 1959.

[3] Straelen, R. A. van, en P. H. Virenque, *De input-output analyse.* Universitaire Boekhandel Uystpruyst, Leuven. 1961.

[4] Venekamp, P. E., *Een nieuw hulpmiddel voor gemeentelijke economische politiek.* Openbare Les. De Erven F. Bohn N.V., Haarlem. 1957.

[5] Netherlands Economic Institute, *Input-Output Relations.* Proceedings of a Conference on Inter-Industrial Relations Held at Driebergen, Holland, 1953. H. E. Stenfert Kroese N.V., Leiden. 1953.

[6] Centraal Bureau voor de Statistiek, *De productiestructuur van de Nederlandse Volkshuishouding.* Twee delen. Uitgeversmaatschappij W. de Haan N.V., Zeist. 1960.

4. ECONOMETRISCHE MACRO-MODELLEN

1. NOGMAALS: SAMENHANG IN DE ECONOMIE

We hebben ons in het vorige hoofdstuk bezig gehouden met het onderwerp van samenhang in het economisch leven, maar toch eigenlijk op een heel beperkte manier. Er is meer dan alleen levering van het ene bedrijf aan het andere en aan de consument. We hebben in de practijk ook te maken met problemen van werkgelegenheid en werkloosheid; zonder twijfel zijn die nauw verwant aan het niveau waarop het bedrijfsleven produceert, maar het vraagstuk van de werkloosheid (d.w.z. hoe die te vermijden) is toch wel zo belangrijk, dat het raadzaam is daar afzonderlijk aandacht aan te besteden. We hebben ook te maken met betalingsbalansproblemen. Stagneert de export vergeleken met de import, dan moet daar iets aan worden gedaan. Ook verdelingsproblemen zijn van groot belang. Daalt het aandeel van de lonen in het nationale inkomen, dan voelen de vakbonden zich geroepen er bij de Regering op aan te dringen om maatregelen te nemen die tot een (naar hun smaak) betere situatie zullen leiden.

De hier genoemde problemen zijn voorbeelden van de onderwerpen waar de zgn. macro-economie zich mee bezig houdt. Kort gezegd, dit onderdeel van de economie richt zich op de grote lijnen van het economisch proces in een land als geheel. Dus niet op detailproblemen zoals de afzet van melkpoeder of de werkloosheid in Zuidoost-Drente (hoe belangrijk die ook voor de betrokkenen mogen zijn). Neen, het gaat om begrippen als het nationale inkomen, de totale werkgelegenheid in Nederland, de totale investeringen en de totale consumptieve bestedingen, het prijsniveau van alle investeringsgoederen gezamenlijk en van alle consumptiegoederen gezamenlijk. Wat is de reden van deze belangstelling voor zulke globale 'aggregaten'? Gedeeltelijk natuurlijk wetenschappelijke nieuwsgierigheid. Door de onderlinge relaties na te gaan kan men proberen het economisch proces te doorgronden. Uiteraard is dat in beginsel evenzeer mogelijk door meer gedetailleerd te werk te gaan; het zou zelfs het voordeel hebben dat ons inzicht daardoor minder grof, meer gedetailleerd wordt. Maar het reali-

seren van dit beginsel, indien consequent doorgevoerd, is onbegonnen werk. Men kan er eenvoudig niet aan denken rekening te houden met de prijs die Mevrouw Jansen bij banketbakker Pietersen voor roomsoezen betaalt wanneer het gaat om het economisch proces in Nederland als geheel.

Naast de wetenschappelijke nieuwsgierigheid is er nog een meer practische reden voor de belangstelling voor hetgeen de macro-economie presteert. Uit de voorbeelden van de eerste alinea (werkloosheid, betalingsbalans, aandeel van de lonen in het nationale inkomen) blijkt wel, dat het aanpakken van de daar genoemde problemen niet zozeer een zaak is van afzonderlijke personen of gezinnen of bedrijven maar eerder op de weg ligt van de overheid. Voor zover de macro-economie practisch werk verzet is dat in de eerste plaats wat men 'overheidseconomie' zou kunnen noemen. Deze term moet niet misverstaan worden; het gaat niet om de interne economische handelingen van overheidsorganen (zoals papier kopen voor de Departementen), ook niet om de huishoudingen van overheidsbedrijven (gas, water, licht, P.T.T., enz.), althans niet in de eerste plaats. Het gaat erom, dat wij eraan gewend zijn aan de overheid een belangrijke taak toe te kennen t.a.v. het landelijke economische proces als zodanig. Stijgt bijv. de werkloosheid boven een niveau dat we acceptabel vinden, dan vinden we vrijwel automatisch ook dat de overheid daar iets aan moet doen; het is uit de tijd te menen dat de slachtoffers maar door particuliere weldadigheid geholpen moeten worden. Vroeger was dat anders; toen stond de rechtspraak en de zorg voor de gevangenissen centraal, de rest was minder belangrijk. Wij willen niet beweren dat rechtspraak en gevangenissen nu onbelangrijk zijn, wel dat hun positie relatief in betekenis is verminderd.

Bezien wij nog eens het voorbeeld van de klimmende werkloosheid. Stel dat hij stijgt boven zeker niveau en dat men van mening is dat hij niet binnen afzienbare tijd vanzelf tot een acceptabel niveau zal terugkeren. Wat kan de overheid daaraan doen? Het eenvoudigst is de werklozen in overheidsdienst aan te stellen; ze kunnen bijv. graafwerk verrichten voor wegen, dijken, viaducten, enz. Maar dit is kennelijk niet aantrekkelijk, al was het alleen maar omdat de meesten daar niet voor opgeleid zijn. Werkloosheid is het gevolg van stagnatie van de economische ontwikkeling; wat de overheid moet doen is die ont-

wikkeling weer op gang helpen, niet door het gebruik van lapmiddelen die stagnatie aan het gezicht onttrekken. Dit kan bijv. door belastingverlaging. Worden de tarieven van loon- en inkomstenbelasting gereduceerd, dan stijgt daardoor het inkomen dat de consument na aftrek van belasting overhoudt (zijn 'beschikbare inkomen' stijgt). Het ligt in de rede dat hij dan meer zal besteden. Dat leidt tot grotere omzetten van de winkeliers, die op hun beurt weer meer bij de groothandel zullen bestellen, bij de fabrieken, kortom, het gaat van hand tot hand en daardoor zijn ook weer meer handen noodzakelijk om de grotere orders te verzorgen. Belastingverlaging leidt dus, naar men mag aannemen, tot geringere werkloosheid; en daar was het om te doen. Maar we zijn er nog niet, want zo'n overheidsmaatregel kan zekere neveneffecten hebben. Zo ook hier, bijv. wat de import betreft. Immers, wat de consument voor zijn extra beschikbare inkomen in de winkel koopt zal gedeeltelijk uit het buitenland afkomstig zijn: sinaasappels, auto's, koelkasten, enz. Niet alleen komt dan het voornaamste effect aan het buitenland ten goede (wat op zichzelf, als we wat ruimer denken, zo erg niet is), maar bovendien kan dit leiden tot betalingsbalansmoeilijkheden. Dat behoeft natuurlijk niet zo te zijn; hebben wij voldoende goud en vorderingen in vreemd geld, dan kunnen we het een hele tijd uitzingen. Maar het kan ook anders zijn en onder zulke omstandigheden zal de overheid er de voorkeur aan geven de werkloosheid te reduceren met maatregelen die dergelijke neveneffecten niet hebben.

Het is mogelijk deze situatie nog veel verder uit te spinnen, maar voor ons doel zijn de volgende conclusies voldoende:

(1) Voor het voeren van een verstandig macro-economisch beleid is het van groot belang, dat de consequenties van de overwogen maatregelen kunnen worden overzien. Het gaat hier om het pure feit dàt verlaging van de belastingtarieven leidt tot hogere consumptie en lagere werkloosheid en (eventueel) betalingsbalansmoeilijkheden.

(2) Het 'hoeveel' is evenzeer van belang. Het is voor de Minister van Financiën niet voldoende te besluiten dat de belastingtarieven gereduceerd worden; hij dient ook een beslissing te nemen over de omvang van de reductie. Dat die reductie groter zal moeten zijn naarmate de werkloosheid ernstiger vormen aanneemt, is een voor de hand liggende opmerking, maar hij brengt ons niet veel verder. Het gaat hierom: als de

belastingtarieven met zoveel procent worden verminderd, hoeveel daalt dan de werkloosheid? Pas als die vraag beantwoord wordt is er een concrete basis voor de te nemen maatregel. Kortom, het probleem is *kwantitatief* van aard.

(3) Voorts moet rekening worden gehouden met de samenhang van de verschillende onderdelen van het economisch proces. We zagen al, dat de belastingmaatregelen van invloed zijn èn op de consumptie èn op de werkloosheid èn op de betalingsbalans. Maar dat is niet meer dan het begin. Stijgt de consumptie, dan zullen i.h.a. de winsten ook stijgen, hetgeen kan leiden tot een grotere bereidheid van ondernemers om in gebouwen machines, enz. te investeren. Kortom, de verdere doorwerking van de maatregelen is een zaak die niet uit het oog mag worden verloren. Ook hier zijn we natuurlijk sterk geïnteresseerd in de mate van doorwerking, dus het kwantitatieve aspect.

(4) Ook de *tijd* vormt een aspect dat niet verwaarloosd mag worden. In het algemeen zijn maatregelen niet onmiddellijk effectief; er gaat een zekere tijd over heen soms een paar dagen, soms veel langer, voor men iets van de gevolgen merkt. Bij de verdere doorwerkingen spelen dit soort 'vertragingen' een nog belangrijker rol. En ook hier is de vraagstelling kwantitatief: na hoe lang hoeveel?

De realisatie vau deze verlanglijst is uiteraard niet eenvoudig en de lezer dient niet van ons te verwachten dat dit hoofdstuk de Sinterklaas zal zijn. Exacte uitspraken, 'foolproof', op al die vragen tegelijkertijd zijn in het huidige stadium niet te geven. Ons voorstel is met wat minder genoegen te nemen, nl. met uitspraken die weliswaar duidelijk streven in de hier aangegeven richting, maar die anderzijds niet meer pretenderen dan benaderingen te zijn. (Dit is natuurlijk geheel analoog aan wat wij deden bij de aan- en afvoeranalyse van het vorige hoofdstuk, toen wij veronderstellingen maakten waarvan tevoren kon worden gezegd dat ze niet voor 100% waar konden zijn.) Wat wij zullen doen is het opstellen van een aantal wiskundige verbanden een aantal relaties, die elk de werking van een onderdeel van het economisch proces beschrijven. Het geheel van die relaties indien volledig (d.w.z. numeriek) gespecificeerd, wordt een *econometrisch model* van het economisch proces genoemd. Op dit terrein van onderzoek heeft ons land een indrukwek-

kende traditie, sinds de Rotterdamse hoogleraar Tinbergen in de dertiger jaren het eerste model (voor Nederland) construeerde en daarna voor de Verenigde Staten en Engeland; later werd hij gevolgd door talrijke anderen.

Het doel van dit hoofdstuk is te laten zien hoe dit soort modellen er uit ziet. Daarbij zullen wij stap voor stap te werk gaan en beginnen met een tweetal modellen, die wat realisme betreft niet veel pretenties kunnen hebben. Zij stellen ons in staat de begrippen te illustreren die bij de hantering van deze modellen worden gebruikt; en ook, hoe men de doorwerking van overheidsmaatregelen in het verloop van de tijd kan nagaan. Het slot van dit hoofdstuk is gewijd aan het door het Centraal Planbureau gebruikte model, dat aanmerkelijk meer gedetailleerd is. Wat wij hier niet zullen bezien is hoe men deze modellen maakt. d.w.z. hoe men ze construeert op basis van economisch-theoretische overwegingen en statistisch waarnemingsmateriaal. De eerste beginselen van dat onderwerp komen in Hoofdstuk 12 aan de orde.

2. EEN HEEL EENVOUDIG MODEL

Wij zullen ons in eerste instantie bezig houden met een wereldje, waarin alleen de totale consumptie (C), de totale investeringen (I) en het nationale inkomen (Y) van een zeker land een rol spelen. Dat land is afgesneden van de buitenwereld, import en export bestaan dus niet.

Stijgt het inkomen, dan stijgt de consumptie. We nemen aan dat de situatie in dat land zo is. dat steeds 80% van het inkomen wordt geconsumeerd:

$$C = 0{.}8\,Y.$$

Dit is een 'consumptiefunctie'; en wel een functie van een heel eenvoudig type. Het kost weinig moeite om ingewikkelder consumptiefuncties te construeren, maar daarop zullen we onze krachten voorlopig niet beproeven.

Nationaal inkomen en nationaal product zijn identieke begrippen. Dat product moet ergens voor dienen, hetzij voor onmiddellijke consumptie (C), hetzij om in de toekomst meer te kunnen consumeren. d.w.z. door in het heden te investeren (I) in machines gebouwen e.d Naar buiten kan het product niet gaan want er is geen export. Dus:

$$C + I = Y,$$

Daarmee is het model wat de vergelijkingen betreft behandeld. Het is een stelsel van twee vergelijkingen in drie variabelen (C, I en Y). Hoe moeten wij er tegenover staan? Met scepsis natuurlijk – niemand zal geloven dat dit soort primitieve wijsheid in wat voor concrete situatie dan ook ons veel zal helpen. Maar het dient hier niet voor een concrete situatie; het gaat ons in de eerste plaats om een aantal begrippen. We verlaten dan de vergelijkingen en richten ons op de drie variabelen, i.h.b. de investeringen (I). Stel eens dat het zo is, dat die investeringen bepaald worden door psychologische invloeden of door toekomstverwachtingen die niets met enige economische realiteit te maken hebben. Dit betekent, dat de hoogte van de investeringen bepaald wordt door factoren, die geheel los staan van het economisch proces dat door ons model wordt afgebeeld. In dat geval noemen we de investeringen een *exogene variabele* m.b.t. dit model. Dus a.h.w. een variabele van buiten af. Voor de consumptie en het nationale inkomen ligt de zaak anders. Het model beschrijft immers door middel van zijn consumptiefunctie hoe het niveau van de consumptie bepaald wordt; en de opbouw van het inkomen uit consumptie en investeringen ($Y = C + I$) bepaalt dan hoe hoog het inkomen is. We drukken dit uit door te zeggen, dat de consumptie en het nationale inkomen de *endogene variabelen* van het model zijn.

Ons model bestaat dus uit twee vergelijkingen in twee endogene en één exogene variabele. Die exogene variabele wordt van buiten af bepaald; en gegeven de waarde die door die variabele wordt aangenomen is het dan mogelijk om voor de twee endogene variabelen de oplossing te vinden. We hebben dan immers twee lineaire vergelijkingen in twee onbekenden en dus (in de regel) een eenduidige oplossing voor beide. En hiermee hebben we het doel van een econometrisch model aangegeven: hoe groot of hoe klein het ook is, het gaat er steeds om het economisch proces af te beelden op een zodanige manier, dat de endogene variabelen rechtstreeks of indirect door de exogene variabelen bepaald worden. Een eenduidige bepaling is natuurlijk alleen mogelijk, wanneer het aantal vergelijkingen gelijk is aan het aantal endogene variabelen. Aan deze eis is hier voldaan (beide aantallen zijn 2). In dat geval spreken we van een *volledig* model en het is tot deze categorie dat wij ons zullen bepalen in hetgeen volgt.

Zojuist werd gesproken over endogene variabelen die 'recht-

streeks of indirect' door exogene variabelen worden bepaald. En inderdaad, 'indirect' is hier zeker op zijn plaats. De consumptiefunctie drukt de endogene consumptie niet uit in de exogene investeringen, maar in het evenzeer endogene nationale inkomen. De andere vergelijking drukt het endogene nationale inkomen maar gedeeltelijk uit in de exogene investeringen; ook de consumptie speelt daar immers een rol. Het is echter niet moeilijk hierin verandering te brengen; we kunnen nl. op eenvoudige wijze elk van de endogene variabelen expliciet in de exogene uitdrukken. Daartoe substitueren we de consumptiefunctie in de vergelijking voor de opbouw van het nationale inkomen:

$0{,}8Y + I = Y$ of ook: $I = 0{,}2Y$,

waaruit onmiddellijk volgt:

$Y = 5I$.

En aangezien de consumptiefunctie zegt, dat de consumptie 80% van het nationale inkomen bedraagt, hebben we dan ook:

$C = 4I$.

Zo vinden we, dat er twee manieren zijn om met hetzelfde model een economisch proces af te beelden. De eerste bestaat (in ons geval) uit een consumptiefunctie en de opbouw van het nationale inkomen. Dit is wat we noemen de *structurele vorm* van het model en de vergelijkingen daarvan heten *structuurvergelijkingen*. De reden is dat die vergelijkingen pretenderen de structuur van een onderdeel van het economisch proces weer te geven; zo geeft de consumptiefunctie weer datgene wat in de consumptieve sector van de economie gebeurt. Daarnaast hebben we het tweetal vergelijkingen $Y = 5I$, $C = 4I$, dat verkregen is door de structuurvergelijkingen zodanig te herleiden, dat elk der endogene variabelen expliciet in de exogene wordt uitgedrukt. Een dergelijk tweetal vergelijkingen wordt dan ook de *herleide vorm* van het model genoemd.

Beide vormen hebben hun nut. De structurele vorm is het handigst wanneer we ons vooral interesseren voor de relaties, de samenhangen, van de diverse onderdelen van het economisch proces. De herleide vorm richt zich daarentegen op het proces als geheel. Hij is het handigst wanneer we de volgende vraag stellen: indien die van buiten af bepaalde exogene variabelen op die-en-die manier veranderen, hoe werkt dat dan door op de endogene variabelen? Voor ons primitieve modelletje is het antwoord eenvoudig genoeg: wat er ook met de investeringen

gebeurt, steeds is de consumptie vier keer zo groot als de investeringen en het nationale inkomen vijf keer.[1]

3. EEN KLEIN MODEL VAN DE VERENIGDE STATEN

We gaan nu de eerste stap in de richting van groter realisme zetten door een iets omvangrijker model te beschouwen, nl. van zes vergelijkingen. Het is afkomstig van de Amerikaanse econometrist L. R. Klein, die het construeerde aan de hand van statistische gegevens van de economie van de Verenigde Staten in de periode tussen de twee Wereldoorlogen. Uiteraard is het feit, dat het model op statistische gegevens berust, een tweede reden waarom we de nu volgende vergelijkingen met iets meer ontzag kunnen bezien.

In de eerste plaats weer de consumptie. Er wordt van uitgegaan, dat het consumptieve gedrag van gezinnen van loontrekkers niet hetzelfde is als dat van de overige gezinnen. De tweede groep bestaat voor een belangrijk deel uit gezinnen wier inkomen uit bedrijfswinst afkomstig is, althans gedeeltelijk, en dit heeft twee effecten. Enerzijds zal deze groep een grotere neiging hebben om van een extra verdiende dollar een deel opzij te leggen om in bedrijf of effecten te beleggen. Dit betekent dat een gegeven extra inkomen leidt tot geringere extra *consumptieve* uitgaven dan bij de loontrekkers het geval is. Anderzijds worden winstinkomens vaak met een zekere vertraging uitbetaald (dus later dan wanneer de winsten in feite gemaakt zijn), zodat de bijbehorende consumptie dan ook pas met zekere vertraging zal reageren. De vergelijking luidt:

(1) $C = 16{,}8 + 0{,}02P + 0{,}23P_{-1} + 0{,}80\ (W_1 + W_2)$

en kan als volgt geïnterpreteerd worden. De totale consumptie (in miljarden dollars per jaar) bevat in de eerste plaats een

1. Het is belangwekkend op te merken, dat de aan- en afvoeranalyse van Hoofdstuk 3 op dezelfde manier te werk gaat. Het model bestaat uit evenveel vergelijkingen als er sectoren zijn, de totale productie per sector is endogeen, de consumptie per sector exogeen. Voor het land-bouw-industrie-dienstensector voorbeeld bestaat het model uit de vergelijkingen op blz. 94 met C_L, C_I resp. C_D in het rechterlid. De daaronder gegeven oplossing is dan de herleide vorm: elk van de endogene productievariabelen is daar expliciet uitgedrukt in de exogene consumptievariabelen.

harde kern van 16,8 miljard[1], in de tweede plaats een stuk dat te maken heeft met winstinkomen (P staat voor 'profits') en in de derde plaats een stuk dat betrekking heeft op looninkomen (W staat voor 'wages'). Wat dit laatste stuk betreft, $W_1 + W_2$ geeft weer het totale looninkomen; W_1 is het looninkomen betaald door het particuliere bedrijfsleven en W_2 het looninkomen verdiend in de overheidssector. Dit onderscheid is gemaakt i.v.m. de derde vergelijking, die hieronder besproken zal worden; voor de consumptie maakt het onderscheid niets uit, want de vergelijking bevat het totale looninkomen $W_1 + W_2$ en discrimineert dus niet tussen de componenten. De coëfficiënt van het looninkomen is 0,80; hetgeen betekent dat iedere extra dollar looninkomen leidt tot een verhoging van de totale consumptie met 80 (dollar) cent. De loontrekkers besparen dus 20% van zo'n extra inkomen.

Nu de winstterm, $0{,}02P + 0{,}23P_{-1}$. Laten we eerst even het verschil tussen P en P_{-1} buiten beschouwing laten, zodat er dan staat $0{,}02P + 0{,}23P = 0{,}25P$. Dit betekent, dat iedere extra dollar winstinkomen leidt tot een verhoging van de consumptie met slechts 25 cent. Er wordt dus inderdaad door deze groep veel meer gespaard uit een gegeven extra inkomen dan bij de loontrekkers het geval is. Maar er staat niet $0{,}25P$; er staat $0{,}02P + 0{,}23P_{-1}$ en die P_{-1} betekent: winstinkomen van een jaar tevoren. Stijgt het winstinkomen van dit jaar met 1 dollar, dan beperkt zich de stijging van de consumptie in dit jaar tot slechts 2 cent. De resterende 23 cent consumptiestijging is pas volgend jaar te merken. Aldus blijkt, dat er inderdaad sprake is van een niet onaanzienlijke vertraging in het consumptieve gedrag van deze groep.

Vervolgens de investeringen. In ons primitieve model hebben we de investeringen als exogeen beschouwd, maar dat is hier niet langer het geval. Het uitgangspunt is, dat stijgende winsten ondernemers optimistischer maken t.a.v. de kans om nieuwe gebouwen en machines rendabel te kunnen gebruiken. Bovendien moeten die investeringen uit zekere middelen gefinancierd worden, en dat wordt duidelijk vergemakkelijkt wanneer de winsten royaal binnenstromen. Vandaar dat de investeringen (I) gerelateerd zijn aan het winstinkomen (P). Een andere fac-

1. Gewerkt is met dollars van constante koopkracht teneinde prijsbewegingen uit te schakelen. Dit is bereikt door de jaarlijkse bedragen te delen door een prijsindex.

tor is de vraag hoeveel gebouwen, machines enz. de ondernemers al hebben staan; kortom, hoe groot de omvang van de bestaande kapitaalgoederenvoorraad is. Het is duidelijk, dat als er al veel is, er naar verhouding weinig animo zal zijn om er nog veel bij te kopen. De investeringsvergelijking luidt dan als volgt:

(2) $I = 17{,}8 + 0{,}23P + 0{,}55P_{-1} - 0{,}15K_{-1}$.

Hierbij staat K voor de kapitaalgoederenvoorraad aan het einde van het jaar. K_{-1} is dus de voorraad aan het einde van het vorige jaar of, wat op hetzelfde neerkomt, aan het begin van dit jaar. De negatieve coëfficiënt (—0,15) laat zien, dat een grotere kapitaalgoederenvoorraad aan het begin van het jaar leidt tot geringere investeringen gedurende het jaar. De coëfficiënt van P (0,23) is belangrijk groter dan de overeenkomstige coëfficiënt (0,02) in de consumptiefunctie. Evenals daar speelt ook hier het vertragingselement een rol; en evenals bij P is ook de coëfficiënt van P_{-1} groter in de investeringsvergelijking dan in de consumptiefunctie (0,55 tegen 0,23). Kennelijk is de winstgevoeligheid van de investeringen groter dan die van de consumptie.

We gaan ons nu bezig houden met het bedrag dat door het bedrijfsleven aan lonen wordt uitbetaald, dus W_1. (De overheidscomponent van de totale loonsom, W_2, wordt exogeen genomen, zoals hieronder nog zal worden uiteengezet.) Wanneer de totale productie van het bedrijfsleven, weer te geven door X, omhoog gaat, zal het noodzakelijk zijn meer personeel in dienst te nemen en daarmee stijgt de loonsom. Omgekeerd zal de loonsom dalen wanneer de productie daalt. Er gaat uiteraard enige tijd over heen voordat tot aanstelling resp. ontslag wordt overgegaan, vandaar dat ook hier met een vertraging wordt gewerkt. De vergelijking ziet er als volgt uit:

(3) $W_1 = 1{,}6 + 0{,}42X + 0{,}16X_{-1} + 0{,}13(t - 1931)$.

De laatste term behoeft nog enige toelichting. In de periode tussen de twee Wereldoorlogen is de machtspositie van de vakbonden in de Verenigde Staten geleidelijk sterker geworden. Dit heeft ertoe geleid, dat de loonsom onderhevig was aan een stijgende beweging (uiteraard afgezien van de invloed van de totale productie, die door de X-termen wordt gerepresenteerd). Die stijgende beweging kwam volgens Klein's berekeningen neer op een jaarlijkse toeneming van 130 miljoen dollar per jaar. Dit dan wordt voorgesteld door de laatste term van de

vergelijking. De variabele t staat voor de tijd, in kalenderjaren gemeten. In 1935 is dus $t - 1931$ gelijk aan 4 en de term in kwestie is $0{,}13 \times 4 = 0{,}52$ miljard dollar. In het jaar daarop, 1936, is $t - 1931 = 5$ en dus is de term $0{,}13 \times 5 = 0{,}65$ miljard dollar, hetgeen inderdaad 130 miljoen hoger is.

De drie tot nu toe besproken vergelijkingen zijn wat men noemt 'gedragsvergelijkingen'. D.w.z. zij beschrijven elk het gedragspatroon van een zekere groep van huishoudingen: het consumptieve gedragspatroon van gezinshuishoudingen en de wijze waarop ondernemers beslissen over hun investeringen en over de omvang van hun personeel. De vergelijkingen die nu volgen zijn meer van het type van boekhoudkundige opsplitsingen die volgens Bartjes moeten kloppen. Zij gaan per definitie op, omdat de variabelen in kwestie volgens zekere regels gedefinieerd zijn; vandaar dat men ze 'definitievergelijkingen' noemt. Een voorbeeld hebben we bij ons primitieve model al gehad, nl. de opbouw van het nationale inkomen uit consumptie en investeringen. Zo hebben we er hier drie. In de eerste plaats geldt, dat de productie van de bedrijven (X) hetzij bestemd is voor consumptie (C) hetzij voor investeringen (I) hetzij voor de overheid. Laten we G schrijven voor de overheidsaankopen bij het bedrijfsleven, dan geldt:[1]

(4) $X = C + I + G.$

In de tweede plaats hebben we de verlies- en winstrekening voor het bedrijfsleven als geheel. Trekken we van de productiewaarde (X) de door het bedrijfsleven betaalde lonen (W_1) af en ook nog de omzetbelasting (T, van 'taxes'), dan krijgen we het winstinkomen:

(5) $P = X - W_1 - T.$

In de derde plaats geldt, dat de kapitaalgoederenvoorraad aan het einde van het jaar (K) gelijk is aan de voorraad aan het begin van het jaar (K_{-1}) plus datgene wat erbij is gekomen door de investeringen (I):

(6) $K = K_{-1} + I.$

1. De productie van het bedrijfsleven kan natuurlijk ook naar het buitenland gaan. Dit is in dit kleine model zo opgelost, dat het saldo van import en export aan de overheidsbestedingen (G) is toegevoegd. Het maakt overigens betrekkelijk weinig uit, omdat de buitenlandse economische contacten van de Verenigde Staten in de periode tussen de twee Wereldoorlogen van geringe betekenis waren vergeleken met de binnenlandse markt.

Hiermee hebben we het model behandeld wat de structuurvergelijkingen betreft. Het zijn er 6 in totaal.

4. NOGMAALS: HET KLEINE MODEL VAN DE VERENIGDE STATEN

We zijn er nog niet. Welke variabelen zijn exogeen, welke endogeen? Is het model volledig? En hoe zit het met die vertragingen? Laten we eerst een lijst van alle variabelen met hun symbolen maken:

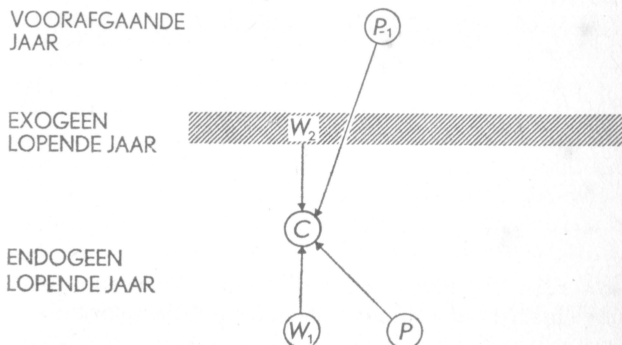
C	consumptie
P	winstinkomen*
I	investeringen
K	kapitaalgoederenvoorraad*
X	productie van het bedrijfsleven*
W_1	loonsom betaald door het bedrijfsleven
W_2	loonsom betaald in de overheidssector
T	belasting te betalen door bedrijven
G	overheidsaankopen bij het bedrijfsleven
t	tijd (kalenderjaren).

Met een sterretje is aangegeven, dat de desbetreffende variabele niet alleen 'lopend' voorkomt maar ook met een jaar vertraging. In totaal zijn er dus 10 variabelen; daarvan komen er 7 alleen maar lopend voor, dus zonder vertraging, en 3 zowel lopend als met een jaar vertraging. Voorbeelden: C komt lopend voor in de vergelijkingen (1) en (4), nergens vertraagd; P komt lopend voor in (1), (2) en (5), en in (1) en (2) bovendien ook nog met een jaar vertraging.

Welke variabelen kiezen we nu als exogeen, dus als van buiten af bepaald? Om te beginnen natuurlijk de tijd (t); het verloop daarvan is onafhankelijk van wat dan ook, dus ook van het economisch proces van de Verenigde Staten dat wij met ons model trachten te beschrijven. Vervolgens bezien wij de drie 'overheidsvariabelen': W_2, T, G. Wil de overheid werkelijk politiek gaan bedrijven – en we hebben in het begin van dit hoofdstuk gezien dat de belangrijkste practische functie van een model is om in dit opzicht een hulpmiddel te zijn –, dan zou het bepaald een zonderlinge zaak zijn wanneer het wel en

wee van die overheidsvariabelen aan het verloop van het economisch proces werd overgelaten. Het gaat er immers om met behulp van die variabelen het proces in de gewenste richting te beïnvloeden! Dit betekent, dat W_2, T en G in beginsel door de overheid bepaald worden, dus niet door het economisch proces in engere zin, dus exogeen zijn. De tijd was al exogeen. Per saldo hebben we dus 4 exogene variabelen, t.w. de onderste vier van de hierboven gegeven lijst van tien. De overige 6 nemen we als endogeen. We hebben ook 6 vergelijkingen, dus is het model volledig.

In verband met de vertragingen is het nuttig enkele tekeningen te bestuderen, zgn. pijlenschema's. Beschouw de volgende figuur:

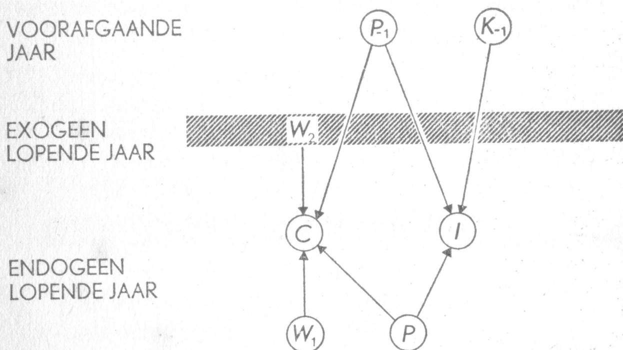


Dit is de consumptiefunctie, vergelijking (1). We tekenen een gearceerde balk, waarin de lopende exogene variabelen genoteerd worden. Hier W_2. Boven de balk noteren we vertraagde variabelen. Hier P_{-1}. Onder de balk noteren we de lopende endogene variabelen. Hier C, W_1 en P. En die verschillende variabelen worden met pijlen verbonden in overeenstemming met de beïnvloeding volgens vergelijking (1), die we hier gemakshalve reproduceren:

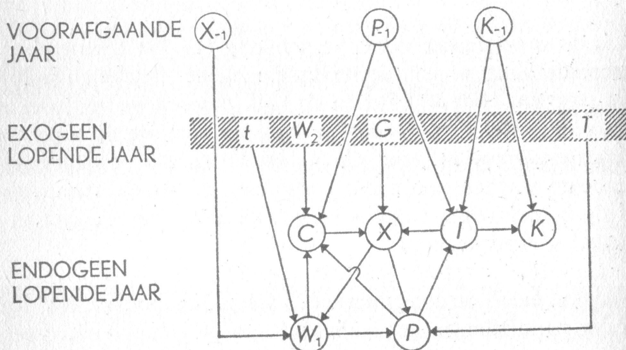
(1) $C = 16{,}8 + 0{,}02P + 0{,}23P_{-1} + 0{,}80\,(W_1 + W_2)$.

Zo komen er bij de consumptie (C) vier pijlen binnen, elk van een der variabelen in het rechterlid van de vergelijking: twee pijlen van andere lopende endogene variabelen (P en W_1), één van een vertraagde endogene variabele (P_{-1}) en één van een exogene variabele (W_2).

Zo gaat men voort met de andere vergelijkingen. Neem bijv. die voor de investeringen, vergelijking (2). Deze bevat uitsluitend endogene variabelen, er is dus geen notering voor een exogene in de balk nodig. De pijlen komen nu bij de investeringen (endogeen, onder de balk) binnen; het zijn er drie in getal, want er staan ook drie variabelen rechts van het gelijkteken in de vergelijking. Eén komt er van een andere endogene variabele (P), twee van vertraagde variabelen (P_{-1} en K_{-1}). We krijgen dan voor consumptie- en investeringsvergelijking tezamen het volgende beeld:



En zo verder. Wanneer alle 6 vergelijkingen hun beurt gehad hebben ontstaat het volgende uiteindelijke pijlenschema:



Wat leert ons dit schema? Het belangrijkste is wel, dat het ons op duidelijke wijze laat zien, hoe de onderlinge relaties door

elkaar heen werken. W_1 beïnvloedt C (de pijl omhoog), C beinvloedt X (de pijl naar rechts) en X beïnvloedt op zijn beurt weer W_1 (de pijl naar links omlaag). Daarmee is deze cyclus van samenhang rond. Maar er zijn meer zulke cycli: P beïnvloedt I, I beïnvloedt X, X beïnvloedt P. Wat we er niet uit kunnen aflezen is de mate van beïnvloeding. Hierin kan worden voorzien door naast de pijlen de desbetreffende coëfficiënten af te drukken (bijv. 0,80 bij de pijl van W_1 naar C), maar daarvan hebben we hier ter wille van de overzichtelijkheid afgezien.

Het pijlenschema leert ons nog iets belangrijks. Om dit te doorzien roepen we in herinnering, dat ons primitieve modelletje geheel 'statisch' was: alle drie variabelen hadden betrekking op hetzelfde tijdvak. Maar het huidige model is 'dynamisch', d.w.z. het bevat variabelen die op *verschillende* tijdvakken betrekking hebben: dit jaar (lopend) en vorig jaar (vertraagd). Daardoor is vrijwel ongemerkt een nieuw onderscheid tussen onze variabelen gekomen. Dit onderscheid maakten we nog niet toen we op blz. 112 een lijst van variabelen opmaakten; maar het wordt wèl gemaakt in het pijlenschema, waar we X, P en K als lopende endogene variabelen onder de balk aantreffen en X_{-1}, P_{-1} en K_{-1} als aparte vertraagde variabelen boven de balk. (In dit model komen geen vertraagde exogene variabelen voor; zijn die er wel, dan noteren we die ook boven de balk, tezamen met de vertraagde endogenen.)

Waar het nu om gaat is dit: het pijlenschema laat zien, dat er eigenlijk maar twee soorten van variabelen zijn. Enerzijds variabelen waar pijlen binnenkomen en ook pijlen vandaan gaan, anderzijds variabelen waar alleen maar pijlen vandaan gaan. De eerste categorie treffen we aan *onder* de balk, dus het zijn de lopende endogene variabelen. De tweede categorie treffen we aan *in* en *boven* de balk, dus het zijn de exogene plus alle vertraagde variabelen. Natuurlijk, het ligt voor de hand dat het zo is: zouden er pijlen binnenkomen bij de exogene variabelen, dan zou dit betekenen dat het economisch proces beschreven wordt met een model dat die variabelen afhankelijk maakt van dit proces, en dan zouden die variabelen niet meer exogeen zijn. En het is ook niet goed denkbaar, dat pijlen bij vertraagde variabelen binnenkomen, want dat zou betekenen dat het verleden bepaald wordt door wat er in het heden gebeurt.

Zo zien we, dat het model zich in feite bezig houdt met de

wijze, waarop de lopende endogene variabelen bepaald worden. Het model specificeert (i) relaties tussen deze variabelen onderling (weergegeven door de pijlen die zowel beginnen als eindigen onder de balk) en (ii) de afhankelijkheid van die lopende endogene variabelen van de rest, dus van de exogene en de vertraagde variabelen (weergegeven door de pijlen die beginnen in of boven de balk en eindigen onder de balk). Uiteindelijk hangen de lopende endogenen helemaal af van de exogenen en de vertraagden; het pijlenschema geeft de onmiddellijke, rechtstreekse samenhangen aan, maar die werken door en het spel wordt in laatste instantie geregeerd door datgene wat het verleden oplevert (de vertraagde variabelen) en wat er van buiten af gebeurt (de exogene variabelen). De wijsgerig getinte lezer zal zich wellicht afvragen: werkelijk geregèèrd? Is er geen plaats meer voor de vrije wil? Ons antwoord is dat zijn reserve gerechtvaardigd is. Maar wij houden ons hier bezig met een model van de werkelijkheid, niet met de werkelijkheid zelf, al hopen wij natuurlijk dat de afwijkingen niet te groot zullen zijn. Die vrije wil zal er heus wel voor zorgen dàt er afwijkingen zijn, maar dat onderwerp stellen wij uit tot het volgende hoofdstuk.

Bij ons primitieve model hoorde een primitieve herleide vorm. Het was alleen maar endogeen en exogeen wat de klok sloeg, niets over lopend en vertraagd. Het pijlenschema heeft dan niets boven de balk, alleen maar in de balk en daaronder. Het zijn dus uitsluitend de (lopende) exogene variabelen die de (lopende) endogene variabelen bepalen en dienovereenkomstig bevat de herleide vorm rechts van de gelijktekens uitsluitend exogene variabelen: $C = 4I$, $Y = 5I$, waarbij I (de investeringen) in dat model exogeen was. Maar nu hebben we alle exogene èn vertraagde variabelen rechts in de herleide vorm, dus alle variabelen die in het pijlenschema vanuit de balk en daarboven hun invloed op het economisch proces doen gelden. Hoe we die herleide vorm afleiden? We hebben 6 lineaire structuurvergelijkingen en 6 lopende endogene variabelen. We moeten naar die variabelen oplossen, d.w.z. ze uitdrukken in alle andere (exogene en vertraagde) variabelen. Doen we dit bijv. voor de consumptie, dan vinden we:

$$C = 41{,}8 + 0{,}74P_{-1} - 0{,}10K_{-1} + 0{,}19X_{-1}$$
$$+ 1{,}34W_2 - 0{,}19T + 0{,}67G + 0{,}16\,(t - 1931).$$

De termen op de onderste regel hebben betrekking op de exo-

gene variabelen, die op de bovenste betreffen de vertraagde variabelen.

Het is bijzonder interessant deze vergelijking te plaatsen naast de consumptiefunctie (1), die we hier nogmaals weergeven:

(1) $C = 16{,}8 + 0{,}02P + 0{,}23P_{-1} + 0{,}80(W_1 + W_2)$.

Beide vergelijkingen hebben links van het gelijkteken de consumptie staan, maar daarmee houdt de verwantschap op. Zou een Minister van Financiën om wat voor reden dan ook de nationale consumptie willen verhogen, zich daarbij *uitsluitend* baserend op de consumptiefunctie (1), dan kunnen we ons voorstellen dat hij bijv. aldus redeneert: 'De vrijheid van het bedrijfsleven is mij lief (binnen zekere grenzen) en ik zal mij daarom niet wagen aan een ingreep wat betreft de winsten (P) en de door de bedrijven uitbetaalde loonsom (W_1); t.a.v. P_{-1} kan ik helemaal niets doen, want dat is een gegeven uit het verleden; en dus, wat mij rest, is een verhoging van de loonsom van de overheidssector (W_2) en blijkbaar stijgt de nationale consumptie met een bedrag gelijk aan 80% van die verhoging.' Maar we weten nu dat we door aldus te redeneren niet verder kijken dan onze neus lang is. Er is meer in de economie dan die consumptierelatie alleen; en wanneer we met de rest ook rekening houden (en ons daartoe op het hier beschreven model baseren), dan moeten we concluderen dat de redenering op twee punten fout is. Ten eerste is de stijging van de nationale consumptie niet 80% maar 134% van de verhoging van de loonsom van de overheidssector. Ten tweede zijn er meer manieren om de consumptie te verhogen, nl. door verlaging van de omzetbelasting (T) of door een verhoging van de overheidsbestedingen bij het bedrijfsleven (G). Waarom? Omdat dit volgt uit de herleide-vormrelatie voor de consumptie, hierboven weergegeven, die in tegenstelling tot de consumptieve structuurvergelijking (1) wèl rekening houdt met de overige samenhangen in het economisch proces.

Voor de overige lopende endogene variabelen gaat men op precies dezelfde manier te werk. Men lost eenvoudig de 6 lineaire structuurvergelijkingen op naar die variabelen en vindt dan voor elk een expliciete uitdrukking in vertraagde en exogene variabelen. Die uitdrukkingen zijn alle lineair en het enige dat daarom van belang is zijn de coëfficiënten. Deze kunnen het eenvoudigst in tabelvorm gerangschikt worden, hetgeen hieronder plaats heeft. De getallen van de eerste rij zijn de coëffici-

enten van de zojuist besproken vergelijking voor de consumptie.

Lopende endogene variabelen	Constante term	Vertraagde en exogene variabelen						
		P_{-1}	K_{-1}	X_{-1}	W_2	T	G	t — 1931
C	41,8	0,74	—0,10	0,19	1,34	—0,19	0,67	0,16
P	38,1	0,86	—0,16	—0,06	0,90	—1,28	1,12	—0,05
I	26,6	0,75	—0,18	—0,01	0,21	—0,30	0,26	—0,01
K	26,6	0,75	0,82	—0,01	0,21	—0,30	0,26	—0,01
X	68,4	1,49	—0,28	0,17	1,54	—0,48	1,93	0,14
W_1	30,3	0,63	—0,12	0,24	0,67	—0,20	0,81	0,20

Zo zien we, dat wanneer de overheid besluit een miljard dollar meer te besteden aan het kopen van goederen bij het bedrijfsleven (m.a.w. G stijgt met 1 miljard), de consumptie C dan stijgt met 0,67 miljard dollar, de winsten P met 1,12 miljard dollar, de investeringen I met 0,26 miljard, enz. Voor de andere overheidsvariabelen (W_2 en T) gaat het op dezelfde manier.

5. OVER GEVOLGEN OP LANGERE TERMIJN

Wij hebben zojuist een Minister van Financiën aan het woord gelaten en hem vervolgens tot betere gedachten gebracht. Stel nu dat hij vertrouwen in ons heeft gekregen en de volgende vraag voorlegt: 'Als ik dit jaar de aankopen bij het bedrijfsleven (G) met een miljard verhoog, dan heb ik nu wel begrepen dat de consumptie (C) in dit jaar daardoor met 0,67 miljard oploopt en de winsten (P) met 1,12 miljard, enz.; maar is daarmee dan de kous af, of gebeurt er *volgend jaar* nog iets? En in het jaar daarna?'

Er gebeurt inderdaad nog iets. Het is nl. een van de kenmerken van een dynamisch model, dat het de consequenties van beleidsmaatregelen als een 'pad in de tijd' laat zien. Op het eerste gezicht lijkt het wellicht alsof er niets gebeurt. De tabel van herleide-vormcoëfficiënten specificeert helemaal niet, dat de lopende endogene variabelen afhankelijk zijn van exogene variabelen met vertraging. Vertraagde exogene variabelen komen immers in dit model niet voor. Maar dat is niet meer dan schijn, want ze spelen impliciet wel degelijk een belangrijke

rol. Laten we aannemen dat onze Minister i.h.b. geïnteresseerd is in de productie van het bedrijfsleven (X); en laten we dan de bijbehorende herleide-vormvergelijking nemen, waarvan de coefficiënten staan in de vijfde regel van de tabel:

$$X = 68{,}4 + 1{,}49P_{-1} - 0{,}28K_{-1} + 0{,}17X_{-1}$$
$$+ 1{,}54W_2 - 0{,}48T + 1{,}93G + 0{,}14(t - 1931).$$

Rechts van het gelijkteken op de eerste regel staan drie vertraagde endogene variabelen: P_{-1}, K_{-1} en X_{-1}. Maar een jaar geleden waren dat *lopende* endogene variabelen! Gaan we ervan uit, dat de structuur van de economie redelijk stabiel is, zodat het model ook op het voorgaande jaar van toepassing was, dan hebben we:

$$P_{-1} = 38{,}1 + 0{,}86P_{-2} - 0{,}16K_{-2} - 0{,}06X_{-2}$$
$$+ 0{,}90(W_2)_{-1} - 1{,}28T_{-1} + 1{,}12G_{-1} - 0{,}05(t - 1932).$$

Dit is niets anders dan de herleide-vormvergelijking voor P, dus de tweede regel van de tabel, maar 1 jaar terugverschoven. Dus wordt P tot P_{-1}, P_{-1} wordt P_{-2}, K_{-1} wordt K_{-2}, enz. Vullen we deze uitdrukking voor P_{-1} in de zojuist gegeven herleide-vormvergelijking voor X in, en doen we hetzelfde voor de overige vertraagde endogene variabelen van de X-vergelijking (dus K_{-1} en X_{-1}), dan krijgen we na enig rekenen het volgende resultaat:

$$X = 129{,}4 + 1{,}33P_{-2} - 0{,}52K_{-2} - 0{,}06X_{-2}$$
$$+ 1{,}54W_2 - 0{,}48T + 1{,}93G$$
$$+ 1{,}54(W_2)_{-1} - 1{,}91T_{-1} + 1{,}93G_{-1},$$

afgezien van een term in t (de tijd), die voor ons doel niet erg belangwekkend is. We zijn er kennelijk in geslaagd om de afhankelijkheid van vertraagde endogene variabelen een jaar verder terug te dringen. Wat we ervoor in de plaats hebben gekregen zijn vertraagde *exogene* variabelen: $(W_2)_{-1}$, T_{-1}, G_{-1}, alle van een jaar geleden. De interpretatie is duidelijk. Worden bijv. de belastingen te betalen door de bedrijven in dit jaar met 1 miljard verhoogd, dan daalt de productie met 0,48 miljard in hetzelfde jaar; maar het jaar daarop daalt de productie met 1,91 miljard! Dat is zo omdat volgend jaar die belastingmaatregel een maatregel van een jaar geleden (T_{-1}) zal zijn. Blijkbaar wordt de Amerikaanse economie door ons model geschetst als een economie met sterk dynamische trekken.[1]

1. Het is puur toeval dat de coëfficiënten van W_2 (lopend en vertraagd) aan elkaar gelijk zijn (1,54). Het is ook niet werkelijk waar. Berekent men nl. een decimaal extra, dan gaan verschillen optreden. Hetzelfde geldt voor de twee coëfficiënten van G.

Zo kunnen we doorgaan. We kunnen P_{-2}, K_{-2} en X_{-2} op dezelfde manier wegwerken, daarna P_{-3}, enz., en het resultaat is dat we het hele pad van de beïnvloeding dan voor ons zien. In beginsel loopt zo'n pad 'in het oneindige' door, maar in de practijk dalen de coëfficiënten na een zeker aantal jaren tot een niveau waarop ze niet meer interessant zijn. Voor de productie van het bedrijfsleven (X) is het pad als volgt:

	W_2	T	G
In hetzelfde jaar	1,54	—0,48	1,93
na 1 jaar	1,54	—1,91	1,93
na 2 jaren	0,99	—1,52	1,24
na 3 jaren	0,34	—0,63	0,42
na 4 jaren	—0,22	0,23	—0,27
na 5 jaren	—0,57	0,82	—0,72
na 6 jaren	—0,72	1,11	—0,90
na 7 jaren	—0,70	1,11	—0,87
na 8 jaren	—0,55	0,91	—0,69
na 9 jaren	—0,35	0,60	—0,44

U ziet, het duurt een hele tijd voordat zo'n maatregel is uitgewerkt. Bovendien slaat het teken om in het vierde jaar: dan hebben de belastingen een positieve invloed op de productie! Bepaald niet een resultaat dat onmiddellijk verwacht zou worden, al moet hieraan worden toegevoegd dat die late positieve invloed van geringere omvang is dan de negatieve invloed op korte termijn. En wellicht overdrijft het model een beetje. Het is nog steeds een klein model met bescheiden pretenties!

6. GROTE MODELLEN

Redelijk serieuze modellen hebben ruim tien tot enkele tientallen vergelijkingen. Het Centraal Planbureau in Den Haag heeft een model van deze omvang in bedrijf; of liever, het heeft sinds jaren een hele reeks van dergelijke modellen gehad. want het maken van een model is geen eenduidige zaak en men verandert waar men op redelijke gronden meent te kunnen verbeteren. De belangrijkste (niet de enige) reden waarom men daar een model hanteert is het Centraal Economisch Plan. Zo'n Plan moet elk jaar gemaakt worden; dat staat in de wet, die op plechtige wijze spreekt van 'een evenwichtig samenstel

van schattingen en richtlijnen met betrekking tot de Nederlandsche volkshuishouding'. U ziet het, dat is in goed-vaderlandse traditie: evenwichtig, geen wilde sprongen. (En inderdaad zullen we in het volgende hoofdstuk zien, dat voorzover van systematische fouten van de voorspellingen van het Plan kan worden gesproken, dit in hoofdzaak de neiging betreft om veranderingen in de nabije toekomst kleiner te ramen dan ze achteraf blijken te zijn.) Voor diegenen van de lezers die principiële bezwaren tegen overheidsplanning hebben kunnen wij onmiddellijk met een geruststelling komen. Het gaat niet om planning in socialistische, laat staan communistische zin. Het pleegt te gaan om de consequenties van alternatieve beleidsmaatregelen van een macro-economisch type. Blijkt de investeringsvergelijking door pijnlijke afwijkingen gekenmerkt te zijn, dan worden geen ondernemers daarvoor ter verantwoording geroepen maar alleen intern de man die de vergelijking heeft gemaakt.

Een relaas over alle (tientallen) vergelijkingen is geen spannend verhaal en we zullen ons daarom beperken tot een enkele. De keus valt op de invoervergelijking, omdat de internationale sector (ook in de aan- en afvoeranalyse van Hoofdstuk 3) tot nu toe stiefmoederlijk is behandeld. Het gaat om de invoer van goederen. niet van diensten. Maar zelfs na aftrek van de diensten is het een bont geheel. Het kunnen goederen zijn die via de importeur onmiddellijk naar de consument gaan, zoals vulpennen en wasmachines. Het kunnen ook halffabrikaten zijn (zoals auto's die hier worden geassembleerd uit ingevoerde onderdelen), maar de belangrijkste component vormen wel de grondstoffen (bijv. granen, die hier verwerkt worden tot brood, beschuit. koek en gebak, en aardolie die verwerkt wordt tot benzine, dieselolie en vele chemische producten). Sommige van de ingevoerde goederen worden verwerkt tot consumptiegoederen, andere worden (na al of niet verder bewerkt te zijn) aan de overheid geleverd, weer andere komen terecht in kapitaalgoederen (machines, schepen, enz.). Ook de exportindustrie eist een deel voor zich op, en tenslotte wordt een deel opgeslagen om pas later verkocht te worden (voorraadvorming).

Zo is het te verklaren dat de invoer van goederen (M) afhankelijk wordt gesteld van de consumptie (C), de overheidsaankopen (G), de investeringen (I), de voorraadvorming (n) en de export, die nog weer wordt onderverdeeld in export van

goederen (E_g) en export van diensten (E_d: vervoer van buitenlanders met de Holland-Amerika Lijn, baggeren in Bombay, enz.). De vergelijking ziet er als volgt uit:

$$M = 0{,}38C + 0{,}39G + 0{,}71I + 0{,}63E_g + 0{,}28E_d + 0{,}79n.$$

De variabelen zijn alle gemeten als veranderingen van jaar tot jaar. Dus staat M voor de verandering van de invoer van goederen (in miljarden guldens), C voor de verandering van de nationale consumptie (idem), enz. De coëfficiënten vertonen onderling vrij aanzienlijke verschillen. Een grote coëfficiënt (0,71) is die van de investeringen. Dit betekent, dat deze bestedingscategorie gekenmerkt is door een groot 'invoergehalte', d.w.z. dat iedere extra gulden die in Nederland geïnvesteerd wordt in gebouwen, machines enz. een belangrijk bedrag aan invoer vereist, nl. 71 cent. De export van goederen (E_g) doet hiervoor niet veel onder: 63 cent. Daarentegen heeft de export van diensten (E_d) een veel geringer invoergehalte. Dat ligt in de rede, want het dienstverleningsbedrijf is arbeidsintensief en vraagt dus naar verhouding minder aanvoer van goederen (uit binnen- of buitenland) en meer mankracht. Voor de voorraadvorming (n) ligt de zaak iets anders i.v.m. het feit dat n niet in miljarden guldens is gemeten. De term $0{,}79n$ betekent, dat een vergroting van de voorraden met 1% aanleiding geeft tot importen ter waarde van 1% van 0,79 miljard, dus 7,9 miljoen gulden.

Dat is dan een van de vergelijkingen. Tezamen met de overige vormt hij een volledig model, waarvoor een herleide vorm kan worden berekend; en daarna kan men, op de wijze waarop dit ook voor het kleine model van de Verenigde Staten is gedaan, het effect van exogene veranderingen op de endogene variabelen aangeven. Ook hier is nl. sprake van een effect in een aantal successieve jaren, omdat het model vertraagde variabelen bevat. Voor een van de versies van het Planbureaumodel zullen wij enkele van deze effecten aangeven.

Eerst het effect van een verhoging van de overheidsuitgaven aan goederen met 1 miljard gulden op het saldo van de betalingsbalans (eveneens in miljarden guldens). Dit is:

in hetzelfde jaar	—0,96
na 1 jaar	—0,43
na 2 jaren	—0,27

De conclusie is, dat wanneer de betalingsbalans er zwak voor

staat, de overheid moet oppassen met het uitgavenbeleid: in het lopende jaar neemt het betalingsbalansoverschot practisch evenveel af (of het tekort vrijwel evenveel toe) als de overheidsuitgaven verhoogd worden en in de twee volgende jaren wordt dit effect nog eens met in totaal 70% verhoogd.

Vervolgens zullen wij nagaan hoe het consumptieprijsniveau (de kosten van levensonderhoud) reageert op veranderingen van het invoerprijspeil en van de loonvoet. Het is duidelijk dat het invoerprijspeil tot stand komt buiten de invloedssfeer van de Nederlandse economie. Het enige dat de overheid op dit punt kan doen is het maken van een raming van de te verwachten ontwikkeling en rekening houden met de consequenties hiervan. Het model dateert uit de tijd van de geleide loonpolitiek, zodat de loonvoet wel degelijk aan overheidsbeslissingen onderworpen was. De cijfers voor de beïnvloeding van de kosten van levensonderhoud luiden als volgt:

	invoer-prijspeil	loonvoet
in hetzelfde jaar	0,24	0,75
na 1 jaar	0,08	0,10
na 2 jaren	—0,01	0,02

In dit geval moeten de coëfficiënten als volgt worden geïnterpreteerd: stijgt het invoerprijspeil in het lopende jaar met 1% t.o.v. het voorafgaande jaar, dan heeft dit een stijging van het consumptieprijsniveau van 0,24% in hetzelfde jaar, van 0,08% in het daaropvolgende jaar, en een daling van 0,01% (dus verwaarloosbaar) na twee jaar tot gevolg. In totaal dus ruim 0,3% in een dr'ejaarsperiode. Loonstijgingen blijken het consumptieprijsniveau veel sterker te beïnvloeden dan stijgingen van het invoerprijspeil: het scheelt ongeveer een factor 3. In beide gevallen is er sprake van een gedeeltelijk vertraagde invloed, maar het dynamische aspect is lang niet zo extreem als wij vonden bij de effecten op lange termijn van het Amerikaanse model.

De tabel die laat zien hoe de bepalende factoren op de endogene variabelen doorwerken heet, in Planbureaujargon, het 'spoorboekje'. Of de vergelijking met de Spoorwegen ooit verder is doorgetrokken is ons onbekend. Maar het zou interessant zijn na te gaan of het bekende devies 'Veilig, Vlug, Voordelig' ook op dit spoorboekje van toepassing is. Wat 'vlug' betreft

kan het antwoord snel gegeven worden. Gaat een Planbureau-ambtenaar naar een vergadering en komt daar een beleidsmaatregel ter sprake, dan heeft hij er maar voor te zorgen het spoorboekje in zijn binnenzak te hebben. Wat hij maximaal moet doen is met potlood en papier een paar kolommetjes op te tellen, nl. wanneer een combinatie van twee of meer maatregelen ter sprake komt. Er bestaat geen concurrerende methode die zo snel zijn antwoord gereed heeft. En 'voordelig' is de methode ook, althans tot op zekere hoogte. Het kost natuurlijk een niet onaanzienlijke hoeveelheid inspanning van hooggekwalificeerde arbeidskracht om een model te maken; echter, is dat er eenmaal, dan is een geduldige stencilmachine voldoende om spoorboekjes te produceren. Maar of de methode ook *veilig* is in de zin dat hij redelijk betrouwbare antwoorden geeft – daarover handelt het volgende hoofdstuk.

LITERATUUR

Tinbergen's eerste model verscheen in een Prae-advies van de Vereniging voor de Staathuishoudkunde en de Statistiek in 1936. Dit is nu minder goed toegankelijk, maar een Engelse vertaling is niet lang geleden verschenen [1, pp. 37-84]. Grotere bekendheid kreeg zijn studie voor de Verenigde Staten [2], in opdracht van de Volkenbond uitgevoerd. Andere Amerikaanse modellen zijn die van Klein [3] en van Klein en Goldberger [4]; de eerstgenoemde studie bevat het model van zes vergelijkingen dat in dit hoofdstuk besproken is. Voor de coëfficiënten van de herleide vorm van dit model alsmede die van de beïnvloeding op langere termijn kan verwezen worden naar Theil en Boot [5]. Twee modellen van het Centraal Planbureau zijn in het kort beschreven in de Aanhangsels van een tweetal Centrale Economische Plannen [6, 7]; de hier besproken invoervergelijking is afkomstig van het eerste model. Voor de beïnvloedingscoefficiënten op blz. 122-123 zie Theil [8].

[1] Tinbergen, J., *Selected Papers*. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam. 1959, pp. 37-84.

[2] Tinbergen, J., *Statistical Testing of Business Cycle Theories*. Twee delen: I. *A Method and Its Application to Investment Activity*. II. *Business Cycles in the United States, 1919-1932*. League of Nations, Geneva. 1939.

[3] Klein, L. R., *Economic Fluctuations in the United States, 1921-1941*. John Wiley and Sons, Inc., New York. 1950.

[4] Klein, L. R., and A. S. Goldberger, *An Econometric Model of the United States, 1929-1952*. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam. 1955.

[5] Theil, H., and J. C. G. Boot, 'The Final Form of Econometric Equation Systems'. *Review of the International Statistical Institute*, Vol. 30 (1962), pp. 136-152.
[6] Centraal Planbureau, *Centraal Economisch Plan 1955*. 's-Gravenhage. 1955.
[7] Centraal Planbureau, *Centraal Economisch Plan 1961*. 's-Gravenhage. 1961.
[8] Theil, H., *Optimal Decision Rules for Government and Industry*. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam. 1964.

5. OVER ECONOMISCHE VOORSPELLINGEN

1. VOORSPELLEN: KUNST OF WETENSCHAP?

De noodzaak van economische voorspellingen kan nauwelijks een punt van discussie zijn. Wie overweegt een maatregel te nemen, doet er goed aan zich te beraden op de consequenties van die maatregel; en veelal ook op de consequenties van alternatieve maatregelen, die wellicht de voorkeur blijken te verdienen wanneer alle gevolgen goed doordacht worden. Soms zijn die gevolgen zo duidelijk dat men daarover niet in het onzekere behoeft te verkeren. Maar dat is lang niet altijd het geval, zodat een voorspelling dan onvermijdelijk is. Zo was er enkele jaren geleden het probleem van werkloosheid onder sigarenmakers. De remedie was een accijnsverlaging op sigaren. Maar welk effect zal een bepaalde accijnsverlaging (bijv. 10%) op de werkgelegenheid van sigarenmakers hebben? Die reductie zal leiden tot een lagere prijs van sigaren; die lagere prijs leidt tot een groter verbruik, omdat de huidige groep van sigarenrokers meer gaat roken en omdat die groep van grotere omvang wordt; en dit grotere verbruik leidt tot een grotere productie van sigaren en dus tot extra vraag op de arbeidsmarkt van sigarenmakers. Maar met hoeveel? Dat staat niet met zekerheid vast; echter, een enigszins betrouwbare voorspelling verdient duidelijk de voorkeur boven volledige onkunde. Daarnaast is het voor de Minister van Financiën van belang het effect van de accijnsverlaging op de fiscale inkomsten te kunnen ramen. Hoe groot de reductie van fiscale inkomsten is hangt weer af van het verhoogde sigarenverbruik, dat het gevolg is van die accijnsverlaging: hoe groter het extra verbruik, des te geringer is de reductie.

Dit voorbeeld valt onder de categorie van zgn. *voorwaardelijke voorspellingen*. Het gaat nl. om de vermeerdering van het sigarenverbruik en van de werkgelegenheid van sigarenmakers onder de voorwaarde, dat de accijns met bijv. 10% wordt verlaagd. In het vervolg van dit hoofdstuk zullen wij nog herhaaldelijk met voorwaardelijke voorspellingen geconfronteerd worden. Wij spreken van *onvoorwaardelijke* voorspellingen indien dergelijke voorwaarden niet worden gesteld. Voorbeeld: Het

nationale inkomen van België zal in 1970 het niveau van 1960 overschrijden met 60%. Voorbeeld van een voorwaardelijke voorspelling: Het Nederlandse nationale inkomen van 1970 zal 65% boven het niveau van 1960 liggen *als* Engeland vóór 1967 tot de Euromarkt toetreedt.

Wanneer dienen we dergelijke voorspellingen au serieux te nemen? Wanneer kunnen we aan een voorspelling het zwaarwichtige adjectief 'wetenschappelijk' toekennen? Of blijft voorspellen steeds een slag in de lucht, een kwestie van 'feeling' of op zijn best 'ervaring'? Welnu, als het een slag in de lucht is (zoals onze Benelux-voorspellingen van zojuist), dan doen we er verstandig aan niet van wetenschappelijk te spreken. Maar we behoeven niet zo pessimistisch te zijn te menen, dat men nog niet boven dit niveau is uitgekomen. Gedurende de laatste decennia, i.h.b. na de Tweede Wereldoorlog, is veel werk verzet om de voorspellingskwaliteit van de economische wetenschap op hoger plan te brengen. Dit ligt voor de hand; goede voorspellingen zijn immers zo niet de enige, dan toch bepaald een zeer belangrijke toets voor hetgeen een vak presteert. Maar wat zijn dan, om met enige hardnekkigheid op de oorspronkelijke vraag terug te keren, de eisen waaraan een wetenschappelijke voorspelling dient te voldoen?

Die eisen vallen alle onder het hoofd *verifieerbaarheid*; maar er is verifieerbaarheid in verschillende opzichten, en daar zullen we enige aandacht aan moeten besteden. Om te beginnen moet het mogelijk zijn te verifiëren of de voorspelling is uitgekomen, zowel in positieve als in negatieve zin. D.w.z., wij eisen dat een voorspelling tot twee uitkomsten kan leiden: hij komt uit of hij komt niet uit, en beide mogelijkheden moeten zich kunnen voordoen. Dit sluit triviale voorspellingen uit. Voorbeeld: Het Nederlandse exportvolume in 1970 zal minder dan het tienvoudige van het niveau van 1963 bedragen. Er is geen schijn van kans dat die voorspelling niet uitkomt, dus een van de twee mogelijke uitkomsten is *niet* mogelijk, dus is de voorspelling triviaal. Ook voorwaardelijke voorspellingen worden uitgesloten als de voorwaarde niet realistisch is. Voorbeeld: Wanneer de Nederlandse industriële productie in 1970 het niveau van 1963 tienvoudig zal overtreffen, wordt het invoervolume 12 maal zo hoog.

Uiteraard vereist verifieerbaarheid óók, dat de voorspelling slaat op een eenduidig gedefinieerd begrip. Toch komt het

soms voor dat aan deze eis niet voldaan wordt. Koning Croesus was er het slachtoffer van. Toen hij een oorlog overwoog en het orakel van Delphi om raad vroeg, kreeg hij ten antwoord: Als Koning Croesus over de Halys trekt, zal hij een groot rijk te gronde richten. Maar er werd niet bij verteld welk rijk het zou zijn. Een ander voorbeeld, van wat meer recente aard: De woningnood in Nederland zal in 1975 ten einde zijn. Dan moet precies worden uitgelegd wat woningnood is. Is die woningnood ten einde zodra ieder gezin en iedere alleenstaande een eigen woning heeft? (Dit vereist een nauwkeurige definitie van het begrip woning; ingewijden weten dat dit geen eenvoudige zaak is.) Maar als de huren hoog zijn zullen er mensen zijn die aan een gedeelte van een woning de voorkeur geven. En er zullen er ook zijn (zoals in feite nu al het geval is), die op een tweede woning beslag leggen: een huisje of een oude boerderij voor de lange weekeinden. Kennelijk dient de voorspeller van de afloop van de woningnood nauwkeurig aan te geven wat hij werkelijk bedoelt.

Onze lijst is nog niet ten einde. De voorspeller dient ook eenduidig aan te geven op welk tijdvak of tijdstip zijn voorspelling betrekking heeft. Zegt hij 'De woningnood zal binnenkort afgelopen zijn', dan hebben we geen mogelijkheid tot verificatie; de voorspeller kan immers steeds volhouden, dat 'zijn tijd' nog niet is aangebroken. Bovendien is het noodzakelijk, dat hij iets naders specificeert over de omvang van de vermoedelijke afwijking tussen voorspelling enerzijds en hetgeen voorspeld wordt anderzijds. In dit verband is het nuttig om een onderscheid te maken tussen zgn. puntvoorspellingen en intervalvoorspellingen. Puntvoorspellingen geven een enkele numerieke waarde. Voorbeeld: Het Nederlandse nationale inkomen in 1970 zal 50,73 miljard gulden bedragen. Een intervalvoorspelling heeft grenzen, binnen welke het voorspelde verschijnsel zal komen te liggen, althans naar de mening van de voorspeller. Voorbeeld: Het Nederlandse nationale inkomen in 1970 zal tussen 49,5 en 52 miljard gulden bedragen. Wat dan de relatie tussen voorspellingen en het voorspelde fenomeen betreft, het is duidelijk dat die in het geval van puntvoorspellingen al heel eenvoudig is. Het zou immers wel bijzonder toevallig zijn, wanneer het nationale inkomen in 1970 na opmeting werkelijk nu net gelijk aan 50,73 miljard gulden zal blijken te zijn. Bij intervalvoorspellingen ligt de zaak anders. Dan zijn er immers

twee reële mogelijkheden: hetzij de realisatie ligt binnen de grenzen (dan heeft de voorspeller gelijk), hetzij hij ligt er buiten (dan heeft hij ongelijk). Kennelijk is de situatie van puntvoorspellingen het moeilijkst. Wat is hun zin, wanneer het van tevoren vaststaat dat zij toch niet uitkomen, dus dat zij met zekerheid of vrijwel met zekerheid tot afwijkingen (voorspellingsfouten) aanleiding zullen geven? Inderdaad, wanneer niets meer bekend is dan dat, kunnen wij weinig wetenschappelijke waarde aan dergelijke voorspellingen toekennen. Maar soms weten we meer. In een aantal gevallen zijn puntvoorspellingen op een dusdanige wijze afgeleid, dat we van tevoren iets kunnen zeggen over de voorspellingsfouten. Dat zijn geen uitspraken met zekerheid, dus niet van het type 'De voorspelling zal 3% te hoog zijn'. In dat geval zouden we die 3% aftrekken en het probleem is van de baan! Neen, het zijn waarschijnlijkheidsuitspraken; en wat verwacht wordt van een wetenschappelijke (dus: verifieerbare) puntvoorspelling is, dat we naast de voorspelling zelf óók een dergelijke uitspraak hebben over de voorspellingsfout. De waarschijnlijkheidsrekening komt in het volgende hoofdstuk aan de orde; in de Epiloog zullen wij op dit punt nog terugkomen.

In de drie voorgaande alinea's ging het in de eerste plaats om de verifieerbaarheid van de voorspelling; maar daarnaast is er ook het probleem van de verifieerbaarheid van de voorspellings*procedure*, dus van de methode die gevolgd is om de voorspelling te maken. Neem bijv. het geval van de puntvoorspelling ad 50,73 miljard voor het Nederlandse nationale inkomen in 1970; stel dat aan alle hierboven opgesomde eisen voldaan is, zodat dus eenduidig vaststaat wat het nationale inkomen precies is (in welke prijzen gemeten, enz.) en er een waarschijnlijkheidsuitspraak is gemaakt over de voorspellingsfout (van welke aard dan ook). Hebben we dan met een wetenschappelijke voorspelling te doen? Niet noodzakelijkerwijs, want het is denkbaar dat die voorspelling het product van pure fantasie is (hetgeen, tussen haakjes, hier het geval is). Het is dus niet voldoende dat de voorspelling zelf verifieerbaar is; de erachter liggende methode, de voorspellingsprocedure, dient het eveneens te zijn. Valt er iets over deze procedures *in het algemeen* te zeggen? Niet bijzonder veel en, naar men mag aannemen, in de toekomst nog minder dan nu, omdat een grote variëteit van methoden het kenmerk is van een vak dat zijn

kinderschoenen is ontwassen. Maar toch wel dit: er is steeds sprake van een tweetal informatiebronnen, nl. enerzijds zekere theoretische overwegingen (hoe simpel die ook mogen zijn) en anderzijds zekere empirische waarnemingen (hoe gering in aantal en onsolide van kwaliteit die ook mogen zijn). Het eenvoudigste dat men kan doen is te voorspellen dat niets verandert. Als het afzetniveau van een verffabriek in 1964 bijv. 50.000 ton bedraagt, houdt men dezelfde hoeveelheid voor 1965 aan; de theorie is dat de afzet constant blijft, de benodigde waarneming is de afzet van 1964. Iets minder eenvoudig (maar toch nog wel heel eenvoudig) is de methode van lineaire extrapolatie; dan 'trekt men het lijntje door'. De theorie is nu dat de afzet*verandering* per jaar constant is, de benodigde waarnemingen zijn die van de afzet in een aantal recente jaren. Een ingewikkelder situatie (waarop we in dit hoofdstuk nog uitvoeriger terugkomen) is de voorspelling met behulp van een econometrisch vergelijkingenstelsel. In dat geval gaat men ervan uit, dat de coëfficiënten van de vergelijkingen constant blijven. Zo ziet men, dat de theoretische overwegingen steeds neerkomen op de veronderstelling, dat iets constant blijft. Dit 'iets' verschilt van geval tot geval naar de aard van de voorspellingsprocedure; naarmate de procedure meer verfijnd is verschuift het 'iets' verder terug van het te voorspellen verschijnsel zelf naar de fundamentele krachten die de bepalende factoren ervan zijn. Uiteraard moet men de veronderstelling van constantheid niet geheel letterlijk opvatten. Het kan voorkomen, dat men goede redenen meent te hebben om bijv. een bepaalde coëfficiënt van een vergelijkingenstelsel te verhogen of te verlagen; men zal dit inderdaad doen en daarmee houdt de coëfficiënt op constant te zijn. Toch wankelt onze conclusie daardoor niet. De wijze waarop men die coëfficiënt zal veranderen, gegeven de redenen die men daartoe heeft, zal nl. aangepast worden aan de wijze waarop men vroeger om soortgelijke redenen soortgelijke coëfficiënten heeft veranderd. Aldus is er toch weer constantheid, zij het 'op hoger niveau' dan dat van de coëfficiënten zelf, nl. wat betreft de wijze van hun aanpassing.

Zo zien we, hoe getracht wordt het voorspellen geleidelijk van kunst tot wetenschap te maken; of liever, hoe gepoogd wordt de objectieve inhoud van de voorspellingsprocedure te vergroten en daarmee het voorspellen een wetenschappelijker

karakter te geven. Geheel kwijt raakt men de subjectiviteit nooit, al was het slechts vanwege de in beginsel nog altijd willekeurige keuze van de te volgen voorspellingsprocedure.

Er is wel eens gezegd, dat voorspellen in de economie gekenmerkt is door een extra moeilijkheid, nl. het feit dat hetgeen voorspeld wordt steeds het resultaat is van handelingen van mensen, elk met een vrije wil. Dat is niet onjuist en het zal een van de oorzaken zijn van de grotere voorspellingsfouten in de economie vergeleken met bijv. de astronomie. Maar het heeft ook voordelen om met mensen in plaats van met planeten en sterren te doen te hebben. Interesseert men zich voor het aantal zonsverduisteringen in de komende tien jaar, dan heeft het weinig zin daarover enquêteformulieren rond te sturen; interesseert men zich echter voor de investeringen door bedrijven in het komende jaar, dan kan men door rechtstreekse ondervraging proberen meer daarover te weten te komen. Natuurlijk, ook dit is niet zonder problemen; de interviewtechniek is een vak op zichzelf. Maar de investeringsplannen van ondernemers vormen een belangrijke bron van informatie om twee redenen, ook al zijn ze meestal geen wetenschappelijke voorspellingen in de zin van hetgeen hierboven werd uiteengezet. Ten eerste vormen die plannen althans gedeeltelijk de blauwdrukken van de werkelijke investeringen in het komende jaar; zij vormen dan een deel van de economie zoals die zich werkelijk ontplooit. Ten tweede is het denkbaar dat zij onderhevig zijn aan systematische afwijkingen t.o.v. de investeringsrealisaties, zodat wetenschappelijke procedures gevonden kunnen worden om die voorspellingen na zekere correcties van betere kwaliteit te doen zijn. Er zijn tegenwoordig talrijke enquêtes van dit soort, ook op het terrein van de consument. Onze eerste concrete toepassing zal een investeringsenquête zijn.

2. DE INVESTERINGSENQUÊTE VAN HET C.B.S.

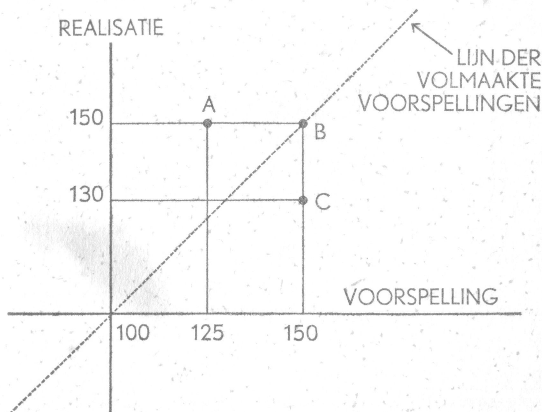
Enkele jaren na de Tweede Wereldoorlog begon het Centraal Bureau voor de Statistiek (C.B.S.) met de organisatie van zijn investeringsenquête. Het gaat daarbij om investeringen door bedrijven in gebouwen, machines, e.d., kortom om investeringen in vaste activa; grondstoffenvoorraden en voorraden gereed product spelen geen rol in de enquête. De gevolgde

procedure komt op het volgende neer. Het C.B.S. zendt enquêteformulieren naar de deelnemende bedrijven (6000 à 8000), waarin geïnformeerd wordt naar de *gerealiseerde* investeringen in het *afgelopen* jaar èn naar de *voorgenomen* investeringen in het *komende* jaar. Per bedrijf dat antwoordt (een zeker percentage pleegt niet te reageren) hebben we dus twee getallen in guldens: een recente realisatie en een plan voor de nabije toekomst. Vervolgens worden alle deelnemende bedrijven gerangschikt in een groot aantal bedrijfstakken en worden investeringsvoorspellingen en -realisaties per bedrijfstak gemaakt. Dit zou een kwestie van eenvoudig optellen zijn wanneer de enquête alle bedrijven van zo'n bedrijfstak besloeg, maar in feite wordt slechts een deel van de bedrijven gedekt. De bedragen per bedrijfstak worden dan geraamd op basis van de veronderstelling, dat het investeringsbedrag per werknemer in de aan de enquête deelnemende bedrijven gelijk is aan het investeringsbedrag per werknemer in de niet deelnemende bedrijven. Aangezien de aantallen werknemers per bedrijf bij het C.B.S. bekend zijn, is het aldus een vrij eenvoudige zaak om tot totale investeringsrealisaties resp. -voorspellingen per bedrijfstak te komen. Uiteraard zit hier een veronderstelling tussen; er is geen enkele reden om aan te nemen, dat de investeringsbedragen voor de twee groepen van bedrijven (gedekt en niet gedekt) steeds precies aan elkaar gelijk zijn. Dit betekent eenvoudig, dat we naast de onvolmaaktheid van de individuele plannen als voorspellingen van de bijbehorende realisaties met nog een tweede bron van afwijkingen te doen hebben, die het gevolg is van onvolledige representatie.

De bedragen per bedrijfstak worden vervolgens opgeteld met als resultaat bedragen voor grotere bedrijfsgroepen: metaalnijverheid, openbare nutsbedrijven, chemische nijverheid, enz. Met deze bedrijfsgroepen (13 in totaal) zullen wij ons bezig houden in hetgeen nu volgt. Samenvattend hebben wij dus per jaar per bedrijfsgroep een voorspelling van de investeringen in het komende jaar en een realisatie over het afgelopen jaar.

Hoe beoordelen we de kwaliteit van die voorspellingen? Wel, door ze te confronteren met de realisatie over hetzelfde jaar, zoals die een jaar later bij de nieuwe enquête ter beschikking komt. Het overzichtelijkst is de beoordeling met behulp van een grafiek. Daartoe meten we de voorspelde en gerealiseerde investeringen per bedrijfsgroep beide als percentage van de rea-

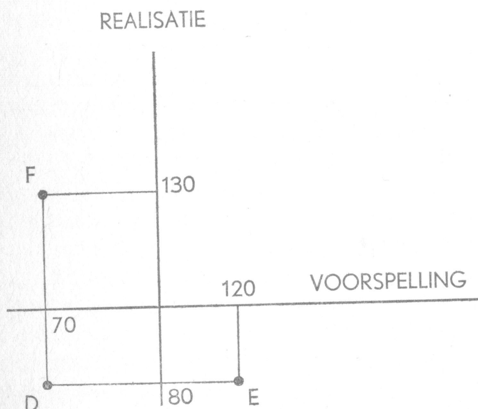
lisatie in het vorige jaar. Bijvoorbeeld, laat de voorspelde investering 50 miljoen gulden zijn en de gerealiseerde (die betrekking heeft op dezelfde periode maar pas een jaar later ter beschikking komt) 60 miljoen; laat voorts het gerealiseerde bedrag van het vorige jaar (dat ter beschikking komt op hetzelfde moment als de zojuist genoemde voorspelling) 40 miljoen zijn. Dan, als percentage uitgedrukt, is de voorspelling 125 en de bijbehorende realisatie 150. In de grafiek zetten we de voorspelling horizontaal af, de realisatie verticaal; het als voorbeeld genomen geval wordt dan gerepresenteerd door het punt *A*.



Was de voorspelling volmaakt geweest, dan was hij niet 125 maar – evenals de realisatie – 150. Dit geval wordt weergegeven door het punt *B*, dat op de gestippelde rechte door de oorsprong loopt. Deze rechte, die een gelijke hoek met de twee assen maakt, is de *lijn der volmaakte voorspellingen*. Voor elk punt op deze lijn is de voorspelling gelijk aan de realisatie en dus een volmaakte treffer. Voor alle punten linksboven de lijn ligt de voorspelling *beneden* de realisatie en wordt dus het investeringsniveau door de voorspelling *onder*schat; dit geldt bijv. voor *A* zoals we zojuist hebben gezien. Voor alle punten rechtsbeneden de lijn der volmaakte voorspellingen ligt de voorspelling *boven* de realisatie en wordt dus het investeringsniveau door de voorspelling *over*schat. Dit geldt bijv. voor het punt *C:* de realisatie is daar 130, de voorspelling 150, dus hoger.

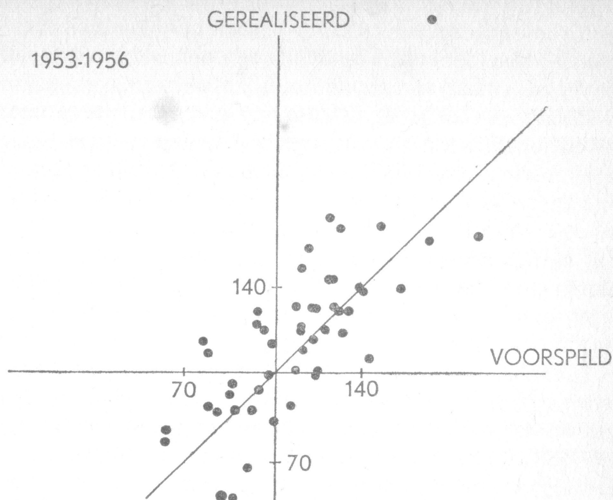
Het is nuttig deze grafische oefeningen nog even voort te zet-

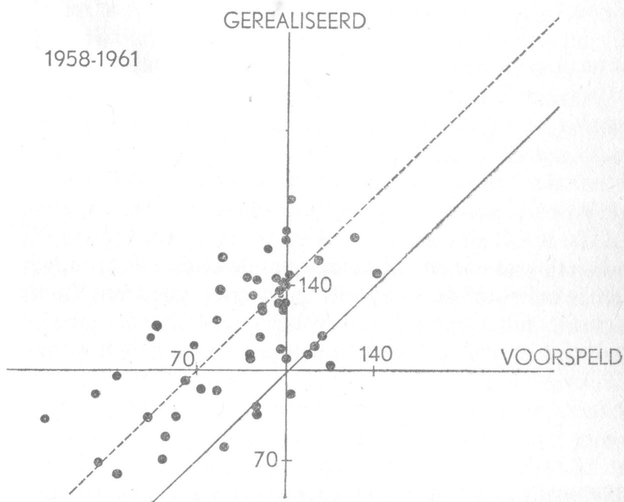
ten. Tot nu toe hebben we ons beperkt tot de gevallen waarin zowel de voorspelling als de realisatie boven 100% liggen, d.w.z. het voorspelde investeringsniveau van het komende jaar ligt boven de realisatie van het jaar tevoren en hetzelfde geldt voor de realisatie van het komende jaar. Kennelijk is dan althans de *richting* goed voorspeld: toeneming voorspeld, toeneming gerealiseerd. Er is nog een andere situatie, waarin de richting ook goed wordt voorspeld: daling voorspeld, daling gerealiseerd. Een voorbeeld is punt *D*, waarvan de realisatie 80 en de voorspelling 70 is.



Zou bij een realisatie van 80 de voorspelling 120 geweest zijn, dan was de richting foutief voorspeld geweest; nl. als stijging terwijl achteraf een daling gerealiseerd blijkt te zijn. We komen dan in punt *E*, rechtsonder de oorsprong. Ook linksboven de oorsprong is de richting foutief voorspeld. Dat toont punt *F*, waar de voorspelling 70 is (dus een daling) en de realisatie 130 (dus een stijging).

Hoe staat het nu in concreto met de investeringsvoorspellingen voor de 13 bedrijfsgroepen? Hierop geven de twee figuren op blz. 135 een antwoord; waarom het er twee zijn zal al heel snel duidelijk worden. De bovenste figuur heeft betrekking op de periode van 1953 t/m 1956; de voorspelling en de bijbehorende realisering per bedrijfstak per jaar worden dus steeds paarsgewijs door een punt weergegeven, zulks op precies dezelfde manier als de punten *A*, *B*, . . . *F* van de figuren op blz. 133-134.


GEREALISEERD
1953-1956
140
70
VOORSPELD
70
140


GEREALISEERD
1958-1961
140
70
140
VOORSPELD
70

(Om tot een wat handzamer figuur te komen is op blz. 135 een logarithmische schaalverdeling gebruikt.) Men ziet dat de punten bepaald niet alle op de lijn van volmaakte voorspellingen liggen, maar aan de andere kant wel een voorspellingskwaliteit van veel hoger gehalte impliceren dan de procedure van extrapolatie op basis van geen-verandering. In dat laatste geval zouden immers alle punten op de verticale as liggen. In feite liggen ze gestrooid rond de lijn der volmaakte voorspellingen, hetgeen uiteraard een veel gunstiger situatie is. Beziet men de figuur nog iets nauwkeuriger, dan valt het op dat de punten vaker linksboven de lijn van volmaakte voorspellingen liggen dan rechtsonder; de eerste categorie omvat ruim 60% van de punten, de tweede nog geen 40%. Dit betekent dat in de beschouwde periode de investeringsvoorspellingen een zekere neiging hadden de realisaties te onderschatten. Beziet men nu de onderste figuur, die betrekking heeft op de periode van 1958 t/m 1961, dan ziet men met een oogopslag dat dit effect nog veel meer geprononceerd is. Bijna 90% van de punten ligt linksboven de lijn der volmaakte voorspellingen. Tal van punten liggen links van de verticale en boven de horizontale as, dus linksboven de oorsprong; in totaal heeft rond de helft van de investeringsvoorspellingen de richting gemist.

Waaraan moeten wij dit verschil tussen de twee perioden toeschrijven? De oorzaak is een verandering van de door het C.B.S. gebruikte definitie van het investeringsbegrip, i.h.b. wat betreft de 'timing' van de investeringen. Vòòr 1957 werd gevraagd naar het bedrag uitgegeven aan de investeringsgoederen die in het desbetreffende jaar werden *geïnstalleerd*. Bij de investeringsplannen: het bedrag dat de ondernemer zich voorstelt te zullen uitgeven aan de investeringsgoederen die in het desbetreffende jaar zullen worden geïnstalleerd. Vanaf 1957 werd gevraagd naar het bedrag uitgegeven aan de investeringsgoederen die in het desbetreffende jaar werden *besteld* (voor plannen analoog). Dit is een belangrijk verschil. Wat het komende jaar geïnstalleerd zal worden is voor een belangrijk deel nu al besteld. Er zullen natuurlijk afwijkingen voorkomen t.o.v. wat in feite geïnstalleerd zal worden, dat zien we ook in de bovenste figuur van blz. 135; maar wanneer het gaat om bestellingen die nog moeten worden gedaan worden allicht tal van toekomstige investeringshandelingen over het hoofd gezien. En dit is ook wat de onderste figuur laat zien: de inves-

teringsvoorspellingen liggen systematisch een stuk te laag. Het is ook duidelijk waarom het C.B.S. tot deze gewijzigde definitie is overgegaan: door naar bestellingen i.p.v. installatie te vragen wordt het mogelijk verder in de toekomst te kijken. Maar even duidelijk is, dat hiervoor een prijs moet worden betaald: in ongeveer de helft van de gevallen wordt een daling voorspeld, die achteraf een stijging blijkt te zijn.

Is het mogelijk hieraan iets te doen? Welnu, een zeer eenvoudige procedure ligt voor de hand. De punten van de onderste figuur liggen kennelijk veel beter rond de gestippelde rechte, die links van de lijn der volmaakte voorspellingen daaraan evenwijdig is getrokken, dan rond de lijn der volmaakte voorspellingen zelf. Die gestippelde rechte snijdt de horizontale as op het punt van een voorspelling gelijk aan 70. Zouden wij alle voorspellingen met 100/70 = 1,43 vermenigvuldigen, dus met 43% verhogen, dan verschuiven wij daarmee alle punten van die figuur evenwijdig aan de horizontale as naar rechts; en wel op een dusdanige manier, dat de punten nu wél rond de lijn der volmaakte voorspellingen komen te liggen. Dat is een aantrekkelijke, eenvoudige correctieprocedure: houd de investeringsdefinitie aan zoals hij nu is, wacht tot de enquêteformulieren binnenkomen en verhoog alle voorspellingsbedragen zonder meer met 43%; de aldus gecorrigeerde voorspellingen zijn beter dan de oorspronkelijke voorspellingen zelf, althans gemiddeld genomen.

Maar werkt deze procedure? Ja en neen. Ja (althans in de hier beschouwde periode), als we naar de individuele bedrijfsgroepen kijken zonder ons te bekommeren om de verschillen in belangrijkheid van die groepen, i.h.b. om de verschillen in investeringsniveaux. Neen, wanneer we met die verschillen wel rekening wensen te houden; de correctie gaat dan nl. veel te ver. We moeten niet uit het oog verliezen, dat de figuren van blz. 135 weliswaar het grote voordeel van aanschouwelijkheid hebben, maar dat zij gebaseerd zijn op voorspelde en gerealiseerde investeringen als percentages uitgedrukt. In de figuren wordt dus in het geheel geen rekening gehouden met de absolute hoogte van de investeringen. En die absolute hoogte is belangrijk; men zal zich nl. in de macro-economische politiek i.h.b. interesseren voor de totale investeringen, die gevonden worden door de investeringen per bedrijfsgroep bij elkaar op te tellen. Stel nu eens dat het zo is, dat de punten van de onderste

figuur van blz. 135 die behoren bij de *grote* bedrijfsgroepen *dicht* bij de lijn der volmaakte voorspellingen liggen; en dat het de middelgrote en kleine bedrijfsgroepen zijn die de zondaars zijn, dus ver naar links liggen. In dat geval betreft de onderschatting van het investeringsniveau in hoofdzaak de kleine investeringsbedragen, zodat de onderschatting van het *totale* investeringsniveau geringer is dan men op grond van het voorgaande zou verwachten.

Of dit inderdaad het geval is valt gemakkelijk genoeg rechtstreeks na te gaan. Het gaat om de vier jaren 1958 t/m 1961; de totale investeringen, zowel voorspeld als gerealiseerd (in miljoenen guldens), alsmede de realisatie uitgedrukt als percentage van de bijbehorende voorspelling volgen hieronder:

	Realisatie	Voorspelling	Realisatie als % van de voorspelling
1958	1990	1799	111
1959	2688	2190	123
1960	3388	2664	127
1961	3771	3073	123

Men ziet, voor het totale investeringsbedrag is de onderschatting lang niet zo groot: het vereiste verhogingspercentage varieert per jaar van ruim 10% tot bijna 30%, dus gemiddeld ruwweg 20%. Dat is de helft minder dan het verhogingspercentage dat de onderste figuur van blz. 135 suggereerde!

Is het dan inderdaad waar, zoals twee alinea's terug werd gesuggereerd, dat grote bedrijfsgroepen aan een relatief geringere onderschatting onderhevig zijn dan kleine? Ook dit kan eenvoudig rechtstreeks worden nagegaan. Hiernaast volgt eenzelfde staatje voor de 13 bedrijfsgroepen, gerangschikt naar afnemend investeringsniveau; de bedragen van de eerste twee kolommen zijn gemiddelden per jaar over de vier jaren 1958—1961 in miljoenen guldens.

Het resultaat laat zien, dat er bij de openbare nutsbedrijven geen en bij de mijnen en de chemische nijverheid nauwelijks sprake is van onderschatting (resp. 98, 103 en 107%). Daarna komt als vierde in de rij van relatief geringe onderschattingen de metaalnijverheid (129%) en als vijfde de sector van de voedingsmiddelen (135%). Maar van die vijf zijn er vier die bovenaan in de lijst naar omvang van de investeringen staan!

	Realisatie	Voorspelling	Realisatie als % van de voorspelling
Metaalnijverheid	703,7	545,6	129
Openbare nuts-bedrijven	614,2	628,0	98
Chemische nijver-heid	466,8	438,1	107
Voedingsmiddelen-industrie	330,1	244,6	135
Textielnijverheid	198,2	133,6	148
Bouwnijverheid	152,6	93,1	164
Papiernijverheid	113,8	76,8	148
Grafische industrie	85,9	49,4	174
Aardewerk, glas, enz.	84,7	59,9	141
Mijnbouw	81,7	79,4	103
Kleding en reiniging	45,1	27,8	162
Leer- en rubber-industrie	44,4	28,0	158
Hout, kurk en stro	38,2	26,6	144

Dus is er inderdaad een duidelijke neiging tot een geringere relatieve onderschatting bij de grootste bedrijfsgroepen. Wenst men correcties op de voorspellingen aan te brengen van het type dat hierboven werd besproken, dan verdient het de voorkeur met dit soort verschillen rekening te houden. Men kan bijv. per bedrijfsgroep verschillende verhogingspercentages toepassen, maar het verdient wel aanbeveling tevoren te verifiëren, of de verhoudingen over de tijd wel voldoende stabiel zijn.

3. ECONOMETRISCHE MODELVOORSPELLINGEN

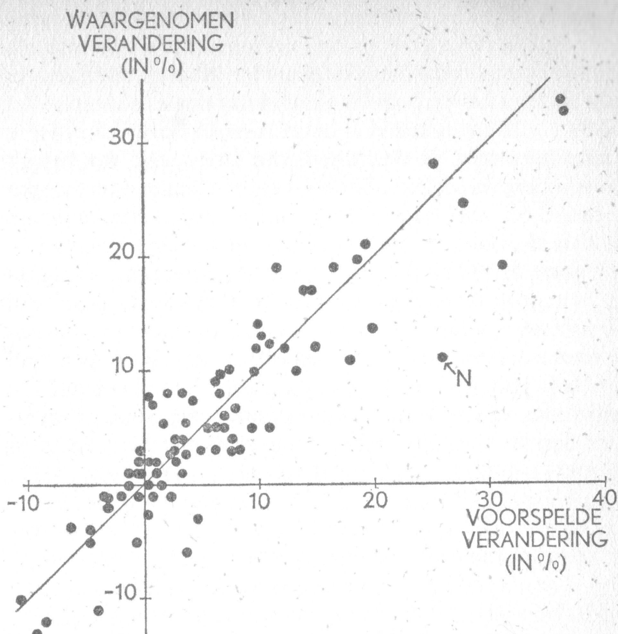
Vervolgens een categorie van voorspellingen, die het wetenschappelijk ideaal in de hierboven aangegeven zin op meer rechtstreekse wijze trachten te benaderen: die van de econometrische macro-modellen. Wij brengen in herinnering, dat de functie van dergelijke modellen neerkomt op de beschrijving van het proces, waardoor zekere (lopende) endogene variabelen worden bepaald. De in aanmerking genomen factoren zijn

daarbij enerzijds de exogene, anderzijds de vertraagde variabelen. Hoe gaat men na of het model aan deze bedoeling beantwoordt? In beginsel is dat heel eenvoudig: door per jaar na te gaan welke waarden die bepalende variabelen (exogeen en vertraagd) in feite hebben aangenomen, door deze waarden in het model in te vullen en dan maar uit te rekenen wat er voor de endogene variabelen uitkomt. Dat zijn dan voorwaardelijke voorspellingen van de door de endogenen aan te nemen waarden; ze zijn voorwaardelijk omdat ze zijn afgeleid onder de voorwaarde, dat de exogenen en de vertraagden zijn zoals ze zijn, d.w.z de waarden aannemen die feitelijk zijn geconstateerd. Is het model van redelijke kwaliteit, dan zullen die voorwaardelijke endogene voorspellingen ook redelijk dicht bij de bijbehorende endogene realisaties dienen te liggen.

Een dergelijk onderzoek is voor een van de modellen van het Centraal Planbureau uitgevoerd door J. Lips en D. B. J. Schouten. Het is niet erg recent; de onderzochte periode betreft de zes jaren van 1949 t/m 1954. De endogene variabelen zijn 14 in getal[1] en het totaal aantal waarnemingen is dus $6 \times 14 = 84$. Voor elk van deze 84 gevallen is berekend de feitelijke procentuele verandering sinds het vorige jaar alsmede de procentuele verandering die geïmpliceerd wordt door de voorwaardelijke voorspelling. De figuur, die is afgedrukt op de volgende bladzijde bovenaan, verschaft een duidelijk overzicht van het geheel. De punten liggen betrekkelijk dicht bij de lijn van volmaakte voorspellingen en het aantal gevallen van een foutief voorspelde richting is gering. Vergelijken we dit beeld met dat van de investeringsvoorspellingen op blz. 135, dan komen de huidige voorspellingen er bepaald niet slecht af. Wellicht ten overvloede merken we op, dat deze uitspraak niet alle voorspellingen *afzonderlijk* betreft. Neem bijv. het punt, dat in de figuur met N is aangegeven. Dit betreft de invoer van goederen in 1949. De voorspelling is een stijging van 26% t.o.v. het niveau van 1948, de realisatie is een stijging van 11%, dus aanzienlijk minder. Deze voorspelling is dus bepaald niet zo fraai. Maar het gaat hier om het beeld als geheel.

De eerlijkheid gebiedt ons te vermelden, dat het voorspellingsbeeld niet altijd zo fraai is. Zo is er een aantal jaren gele-

1. De volledige lijst van deze variabelen vindt men in de literatuuropgaaf aan het einde van dit hoofdstuk.



den in de Verenigde Staten een onderzoek uitgevoerd naar de voorspellingskwaliteit van een Amerikaans model, dat aangepast was aan statistische gegevens van de periode tussen de twee Wereldoorlogen. Het werd vervolgens toegepast op een naoorlogs jaar, met ietwat povere resultaten. Het verschil in uitkomst kan aan tal van oorzaken liggen: de structuur van een economie kan na een oorlog gewijzigd zijn, de Amerikaanse economie is de Nederlandse niet en kan econometrisch minder gemakkelijk te voorspellen zijn, enz. Het zou te ver voeren hierop in detail in te gaan, maar voor één aspect willen wij een uitzondering maken. Het gaat hier om voorwaardelijke voorspellingen, waarbij i.h.b. de exogene variabelen als gegeven worden aangenomen. En het punt is dan dit: hoe belangrijk zijn die exogene variabelen in het economisch proces dat door het model wordt uitgebeeld? Neem bijv. de export. (In het Planbureaumodel is er weliswaar een exportvergelijking, zodat de export strikt genomen endogeen is, maar de jaarlijkse ex-

portveranderingen bevatten een dusdanig omvangrijke exogene component, dat de export in het hier vermelde onderzoek als exogeen is beschouwd.) Nu is de export in Nederland bijzonder belangrijk, nl. ruwweg de helft van het nationale inkomen. Het spreekt vanzelf, dat wanneer zo'n belangrijke factor als gegeven wordt aangenomen, we al een heel eind op weg zijn wat betreft de voorspelling van de rest. In de Verenigde Staten daarentegen heeft de export naar verhouding veel minder te betekenen en dus is het gegeven zijn van een dergelijke variabele en veel geringere waarde voor de voorspelling van de endogene variabelen. Dit geldt algemener: in een klein land als Nederland met zijn relatief belangrijke internationale contacten is nu eenmaal meer exogeen, dus van buiten af bepaald, dan in een land als de Verenigde Staten. Niet dat dit laatste land minder belangrijke internationale economische contacten heeft. Integendeel, ze zijn belangrijker; maar vergeleken met hetgeen er in het binnenland omgaat hebben ze minder te betekenen.

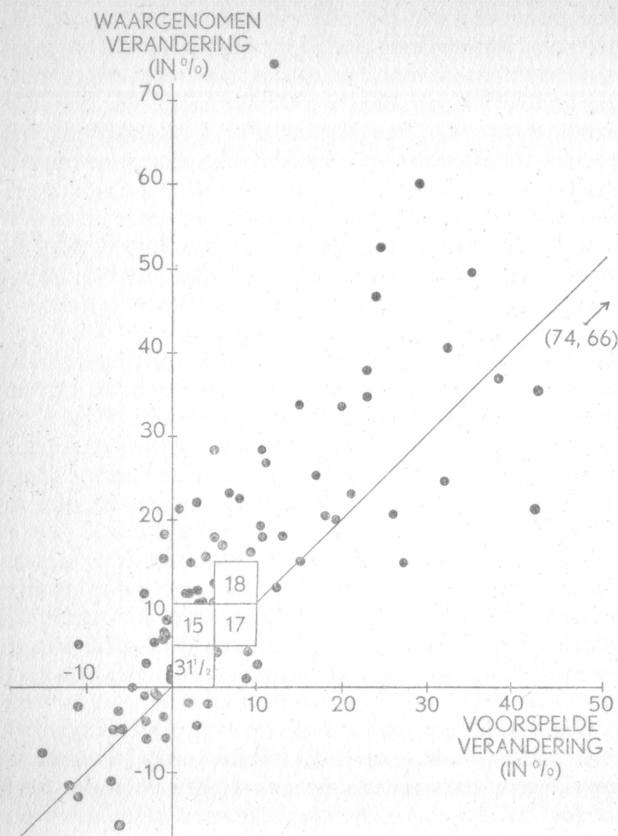
4. DE VOORSPELLINGSKWALITEIT VAN DE CENTRALE ECONOMISCHE PLANNEN

We gaan nu het terrein iets verbreden. Naast de eigenlijke modelvoorspellingen hebben we in de practijk nog met een tweede groep van voorspellingen te doen, nl. die van de exogene variabelen. We hebben tot nu toe de correcte (waargenomen) waarden van die variabelen ingevuld, maar het mag niet uit het oog verloren worden dat men die van tevoren niet kent; bijv. de verandering van het prijspeil van ingevoerde goederen in het komende jaar. Ook die variabelen moeten dus voorspeld worden, waarna deze voorspellingen in het model gesubstitueerd worden om (onvoorwaardelijke) voorspellingen van de endogene variabelen op te leveren. Men zal aanvoelen, dat het voorspellingsbeeld dan over het algemeen van geringere kwaliteit zal zijn; er is immers sprake van twee foutenmogelijkheden i.p.v. één. En dat is de situatie waarmee we in de realiteit geconfronteerd worden. Daarop moet echter één kanttekening gemaakt worden. Toen wij spraken over exogene variabelen, legden we de nadruk op buitenlandse factoren. Maar er is nog een andere belangrijke groep van exogenen, nl. de overheidsvariabelen, de instrumenten dus waarmee de

overheid kan trachten zijn invloed op het economisch proces te doen gelden. Die variabelen zijn juist daarom zo belangrijk, omdat naar wij eerder zagen de practische functie van een econometrisch macro-model gelegen is in de toepassing bij beslissingsproblemen van de overheid. Wordt nu een bepaalde maatregel overwogen, dan betekent dit niets anders dan dat men aan een bepaalde exogene variabele een zekere waarde toekent, althans in gedachten. En die waarde is dan automatisch bekend! Voor de berekening van de doorwerking van een *gegeven* beleidsmaatregel gaat het dus om voorwaardelijke voorspellingen van het type van de figuur op blz. 141; en het is prettig te weten dat die het er in de onderzochte periode over het algemeen redelijk hebben afgebracht.

Nu dan de onvoorwaardelijke voorspellingen, dus inclusief alle fouten van de exogene voorspellingen zoals die in feite gemaakt zijn. Dit komt neer op de jaarlijkse voorspellingen van de Centrale Economische Plannen. Op dit terrein is een onderzoek beschikbaar van C. van de Panne, die de voorspellingen en realisaties van 23 variabelen heeft vergeleken; de periode omvat 7 jaren, nl. van 1949 t/m 1956.[1] In totaal dus 23 × 7 = 161 paren van voorspellingen en realisaties. Ook hier gaat het om procentuele veranderingen t.o.v. het niveau van de desbetreffende variabele in het vorige jaar. Deze 161 paren van prognoses en realisaties zijn (op dezelfde wijze als we reeds eerder in dit hoofdstuk hebben gezien) afgebeeld op blz. 144.[2] Onze conclusies zijn tweeërlei. In de eerste plaats zijn de voorspellingen in kwaliteit gereduceerd vergeleken met de figuur op blz. 141. Het lag voor de hand dat dit zou gebeuren, omdat de voorspellingsfouten van de exogene variabelen nu een rol gaan spelen (op blz. 141 waren die fouten gewoon nul); en we zien dat het inderdaad is gebeurd. Aan de andere kant blijft het

1. De volledige lijst van deze variabelen vindt men in de literatuuropgaaf van het einde van dit hoofdstuk. Het jaar 1952 is buiten beschouwing gelaten i.v.m. een afwijkende opstelling van het Plan.
2. Eén punt (een voorspelde stijging van 74% gevolgd door een waargenomen stijging van 66%) valt buiten de tekening. Dit is m.b.v. een pijl aangegeven. De vier getallen (31½, 15, 17 en 18) representeren de aantallen punten die in de bijbehorende vierkante gebieden thuis horen. De aantallen punten zijn te groot om weergave van ieder individueel punt mogelijk te maken. (De ½ van 31½ heeft betrekking op een punt dat op de rand van het vierkante gebied ligt.)



waar, dat er een duidelijke positieve relatie bestaat tussen voorspelde en gerealiseerde veranderingen. Zou voortdurend 'geen verandering' zijn voorspeld, dan was er van die positieve relatie geen sprake geweest. Natuurlijk is 'geen verandering' nogal onrealistisch in de na-oorlogse ontwikkeling van de Nederlandse economie; er was immers sprake van een overwegende stijging van de meeste variabelen, hetgeen tot uiting komt in het feit dat de grote meerderheid van de punten boven de horizontale as ligt. Maar zelfs wanneer men rekening zou houden met die overwegende stijging door bijv. voor alle varia-

belen in alle jaren een constante procentuele toeneming te voorspellen, dan nog is het hier gepresenteerde voorspellingsbeeld beter. Samenvattend kunnen wij dus stellen, dat de hier beschouwde voorspellingen zekere merites hebben maar dat zij anderzijds duidelijk beneden het niveau van de voorwaardelijke endogene voorspellingen staan.

De tweede conclusie is, dat de onvoorwaardelijke voorspellingen een systematische fout te zien geven. Beschouw daartoe de punten, die boven de horizontale as liggen en bovendien rechts van de verticale as; dus de gevallen waarin een stijging is voorspeld en gerealiseerd. Deze punten liggen in meerderheid *boven* de lijn van de volmaakte voorspellingen, hetgeen betekent dat de voorspelde verandering kleiner is dan de gerealiseerde verandering. Korter: *de verandering is onderschat*. Beschouw vervolgens de punten, die beneden de horizontale as liggen en bovendien links van de verticale as; de punten linksonder dus, die de gevallen representeren waarin een daling is voorspeld en gerealiseerd. We vinden hier, dat die punten voor het grootste deel onder de lijn van de volmaakte voorspellingen liggen, zij het dat het effect nu iets minder geprononceerd is. Ook voor deze punten geldt, dat de verandering is onderschat: de voorspelde daling is bijv. 5% en de gerealiseerde daling 10%. In totaal hebben we dus dit beeld: als de voorspelling de goede richting aangeeft, dus voorspelde en gerealiseerde verandering hetzij beide positief hetzij beide negatief, dan geldt in de meerderheid van de gevallen (ruim drie kwart), dat de voorspelde verandering te klein is, dus dat de verandering is onderschat. Overschatting komt voor, maar is eerder uitzondering dan regel. Wanneer we ook nog met de derde mogelijkheid rekening houden, nl. met het geval van een foutief voorspelde richting (de punten linksboven en rechtsonder de assen), dan is de verdeling per jaar als volgt:[1] *zie noot pagina 146*

	1949	1950	1951	1953	1954	1955	1956	Totaal
Verandering onderschat	18½	17½	12	12	17	17½	17	111½
Verandering overschat	3½	4½	8	9	3½	1	5	34½
Richting foutief	1	1	3	2	2½	4½	1	15
Totaal	23	23	23	23	23	23	23	161

In totaal zijn er per jaar steeds 23 gevallen, omdat het onderzoek immers op 23 variabelen betrekking heeft. We zien dus jaar in jaar uit een meerderheid, soms een grote meerderheid, van veranderingen die *onder*schat blijken te zijn; een minderheid, vaak een kleine minderheid, van veranderingen die *over*schat blijken te zijn; en steeds een kleine minderheid van gevallen, waarin de richting verkeerd werd voorspeld.

Wat is de oorzaak van die slagzij in de richting van onderschatting van veranderingen? De onmiddellijke oorzaak kan gemakkelijk genoeg gegeven worden. Het moet liggen aan de voorspelling van de exogene veranderingen. In de figuur op blz. 141 is met volmaakte exogene voorspellingen gewerkt en daar was er geen sprake van een systematisch effect t.a.v. te kleine voorspelde veranderingen. Maar dan vragen wij verder: waaraan ligt het dat die exogene veranderingen in de meerderheid van de gevallen onderschat worden? Om deze vraag te kunnen beantwoorden dienen we ons terrein even te verbreden. De slagzij in de richting van onderschatte veranderingen is nl. een veel algemener verschijnsel. De bevolking van Nederland is sinds de Tweede Wereldoorlog meer in omvang toegenomen dan de experts hadden verwacht. Het aantal auto's dat in Nederland rondrijdt is meer toegenomen dan men had verwacht. Zo kan men een lange lijst van dergelijke gevallen geven, niet alleen op economisch terrein. Sportjournalisten geven elke week een voorspelling van de uitslagen van voetbalwedstrijden van de komende zondag. Daarbij baseren zij zich in sterke mate op de prestaties van de elftallen in de voorgaande wedstrijden. Uiteraard bestaat er gemiddeld genomen een zeker verband tussen wat een elftal de komende zondag presteert en hetgeen het de vorige zondagen gepresteerd heeft; maar dit verband is niet zo strak als het verband waarop de sportjournalisten hun prognoses baseren en dus, door al te sterk de nadruk te leggen op de prestaties in het verleden, onderschatten zij de verandering van het prestatieniveau. Uiteindelijk is deze systematische fout terug te voeren tot de wijze waarop de voorspellingen gemaakt worden. Men gaat er immers vanuit, dat

1. De halven van deze tabel zijn het gevolg van de (schaarse) gevallen van volmaakte voorspelling. Die zijn elk voor de helft ondergebracht bij onderschatting, voor de helft bij overschatting. Daarnaast zijn er nog twee punten op de assen, die elk voor de helft ondergebracht zijn bij 'richting foutief'.

'iets' ongewijzigd blijft. Is dat 'iets' het te voorspellen verschijnsel zelf, dan is er onderschatting van veranderingen in extreme zin: de voorspelde verandering is dan steeds nul. Dringt men dieper door in de fundamentele krachten, die het verschijnsel bepalen, dan gaat de slagzij in de richting van onderschatting geringer worden. Een voorbeeld wordt gevormd door de voorwaardelijke endogene voorspellingen. De genoemde slagzij valt daar niet waar te nemen; en inderdaad, een uitvoerig onderzoek naar de bepalende factoren, geformaliseerd in een econometrisch model, ligt aan die voorspellingen ten grondslag. Maar voor de exogene voorspellingen is dit anders. De zaak ligt daar ook veel moeilijker omdat bijv. buitenlandse exogene variabelen gedeeltelijk bepaald worden door de maatregelen van buitenlandse overheden. Het is geen eenvoudige zaak die een jaar vooruit te voorspellen.

5. AAN- EN AFVOERVOORSPELLINGEN

Als laatste illustratie enige voorspellingsresultaten verkregen met behulp van de aan- en afvoeranalyse. Op dit terrein is een tweetal studies beschikbaar van de hand van C. B. Tilanus en G. Rey.

Het betreft weer voorwaardelijke voorspellingen. De aan- en afvoeranalyse pretendeert immers te kunnen aangeven, hoe groot de totale productie per sector is, gegeven het bedrag dat de consumenten aan de producten van elke sector uitgeven. Het ligt daarom voor de hand om op precies dezelfde manier te werk te gaan als bij de voorwaardelijke endogene voorspellingen volgens de econometrische modellen: nl. de feitelijk geconstateerde consumptiebedragen in de aan- en afvoervergelijkingen in te vullen en dan door te rekenen wat er voor de totale productie per sector uitkomt, waarna deze voorwaardelijk voorspelde totale productie met de bijbehorende waargenomen productie per sector kan worden vergeleken. Wij zullen dit voorgestelde program inderdaad uitvoeren, zij het met een amendement. De totale productie bestaat voor een niet onaanzienlijk deel uit hetgeen rechtstreeks door de consument wordt verbruikt. Nu worden, zoals wij zojuist zagen, die consumptieve bedragen rechtstreeks gesubstitueerd en deze zijn dus foutloos. Het is duidelijk, dat we van de voorspellingen een te

optimistisch beeld krijgen wanneer deze voor een belangrijk deel uit foutloze bedragen bestaan. Neem als voorbeeld een sector waarvan we de totale productie in zeker jaar op 100 stellen en de consumptie op 80. Dus is het totaal van onderlinge leveringen, van deze sector afkomstig, gelijk 20. Laat die totale productie als 102 voorspeld zijn, dus een fout van slechts 2%. Maar van die 102 is er een component 80, die er van meet af aan is ingestopt. Het verschil, dus 22, is de voorspelde onderlinge levering. Dat is uiteindelijk het enige dat wèrkelijk voorspeld wordt en de voorspellingsfout (t.o.v. de realisatie 20) is hier 10%. Kennelijk is dit laatste een veel realistischer zienswijze. De remedie is dus eenvoudig: we beschouwen de aan- en afvoeranalyse als een voorspellingsmethode voor hetgeen elke sector aan alle sectoren levert (dus het totaal der onderlinge leveringen per sector), gegeven de bedragen die de consument aan de producten van de diverse sectoren besteedt.

Daarmee is het terrein van onderzoek afgebakend. Vervolgens het onderzoek zelf. Dit wordt gekenmerkt door een practische moeilijkheid, waarmee veel econometrisch onderzoek van de laatste jaren in klimmende mate te maken heeft, nl. die van een mer à boire van cijfers. We hebben aan- en afvoertabellen voor elk van de jaren 1948 t/m 1957, dus 10 stuks. We kunnen *elk* van deze tabellen gebruiken om voor *elk* van de sectoren het totaalbedrag aan onderlinge leveringen te voorspellen voor *elk* van al die jaren. Om precies te zijn: we zullen met 27 sectoren werken; niet met alle 35, omdat voor een aantal de onderlinge leveringen klein zijn en dus niet interessant.[1] Voorts zullen we ons bepalen tot 'vooruit' voorspellen, omdat dit in de practijk het meest interessant is. D.w.z., we zullen de tabel van 1950 gebruiken voor de voorspelling van 1951 en later, niet voor de jaren vóór 1950. Hoeveel voorspellingen (met alle 10 tabellen voor alle 27 sectoren en alle jaren na het jaar van de tabel) zijn er dan in totaal? Na enig rekenen vindt men 1485 stuks.[2] Men kan al die uitkomsten in een grote tabel rangschikken, maar dat wordt een onoverzichtelijke massa. De enige manier om er iets verstandigs van te zeggen is ze groepsgewijs

1. De buiten beschouwing gelaten sectoren zijn de nummers 23, 26, 27, 30 en 32 t/m 35 van de lijst op blz. 82.
2. Rekening houdend met het feit, dat voor 1958 ook al productie en consumptie per sector beschikbaar zijn, maar nog niet een volledige tabel.

te beschouwen en per groep zekere kengetallen te berekenen.

Het is nuttig met een voorbeeld te beginnen. Neem sector 13 (chemische nijverheid) en neem verder de eerste tabel, dus die van 1948. Door de consumptie per sector van 1949 in de bij die tabel behorende aan- en afvoervergelijkingen in te vullen krijgen we voor elke sector: de voorwaardelijke voorspelling van de totale onderlinge leveringen, gegeven het consumptiepatroon van 1949 en op basis van de tabel van 1948. Voor de hier beschouwde sector vinden we dan een voorspelling die, indien geconfronteerd met de in feite waargenomen leveringen van die sector in 1949, 1,5% te hoog blijkt te zijn. Wij noteren dus een voorspellingsfout van 1,5%. Dit herhalen we voor de consumptie en de onderlinge leveringen van 1950, weer met dezelfde tabel van 1948 en voor dezelfde sector. We vinden dan een voorspellingsfout van 1,4%. En zo gaan we door: met de tabel van 1948 berekenen we de voorspelde totale onderlinge leveringen uitgaande van deze sector voor elk van de successieve jaren tot 1958 toe. Het resultaat is de volgende rij van voorspellingsfouten (in %):

1949	1950	1951	1952	1953	1954	1955
1,5	1,4	—11,2	—11,5	—10,7	—7,6	—9,0
1956	1957	1958				
—8,3	—9,7	—11,4				

waarbij de mintekens aangeven, dat de voorspelling beneden de realisatie ligt.

Wat leert ons dit resultaat? Bijzonder regelmatig is het niet, maar toch valt wel te onderkennen dat de voorspellingsfouten in de allereerste jaren minder groot zijn dan in de latere jaren. Dat ligt ook in de rede. We werken hier met de tabel van 1948, die de productiestructuur van dat jaar weergeeft. Die structuur is niet constant maar aan geleidelijke verschuivingen onderhevig. Voor het volgende jaar, 1949, zal dit i.h.a. nog niet zo belangrijk zijn; maar die verschuivingen zullen omvangrijker zijn, althans gemiddeld genomen omvangrijker, naarmate we het jaar van de tabel verder achter ons laten liggen. En dienovereenkomstig hadden we ook redelijkerwijs mogen verwachten, dat de voorspellingsfouten in betekenis toenemen naarmate we latere jaren voorspellen.

Het ligt daarom voor de hand een onderscheid te maken

tussen voorspellingen één jaar vooruit, voorspellingen twee jaren vooruit, enz. 'Eén jaar vooruit' betekent dus dat we de 1948-tabel gebruiken om de onderlinge leveringen van 1949 te voorspellen, of ook de 1949-tabel om 1950 te voorspellen, enz. 'Twee jaren vooruit' betekent dat we de 1948-tabel gebruiken om 1950 te voorspellen of de 1949-tabel om 1951 te voorspellen, enz. Zo kan men doorgaan tot 10 jaren vooruit, hetgeen gegeven het materiaal maar op één manier kan, nl. door met de 1948-tabel de onderlinge leveringen van 1958 te voorspellen. Laten we nu eerst in detail de voorspellingen van sector 13 één jaar vooruit bezien. Dit zijn er 10 in totaal: de voorspellingen met de 1948-tabel voor 1949, die met de 1949-tabel voor 1950, . . ., en tenslotte die met de 1957-tabel voor 1958. De bijbehorende voorspellingsfouten (in %) zijn als volgt:

1,5   0,4   —12,5   —0,7   0,2   3,9   —1,6   0   —2,2   —1,9.

Laten we dan deze groep van voorspellingen als een redelijk homogene groep beschouwen; het gaat immers om voorspellingen, alle één jaar vooruit, van de onderlinge leveringen afkomstig van dezelfde sector. De vraag rijst: kunnen we de kwaliteit van deze voorspellingen door een enkel getal uitdrukken? Dat kan zeker, en wel op talrijke manieren. Men kan bijv. het gemiddelde van de 10 procentuele voorspellingsfouten nemen. Dit heeft echter het nadeel, dat fouten in tegengestelde richting elkaar compenseren. Wordt bijv. een fout van 3% gevolgd door een fout van —3%, dan heffen die elkaar in het gemiddelde precies op; dit tweetal fouten leidt tot precies hetzelfde gemiddelde als we gevonden zouden hebben in het geval van twee volmaakte voorspellingen, hetgeen kennelijk niet redelijk is. Dit bezwaar wordt vermeden door de zgn. gemiddelde kwadratische voorspellingsfout, die verkregen wordt door alle fouten te kwadrateren, op te tellen en te delen door hun aantal. In ons geval:

$$(1{,}5^2 + 0{,}4^2 + \ldots + 1{,}9^2) / 10 = 19.$$

De gedachtengang, die aan de gemiddelde kwadratische voorspellingsfout ten grondslag ligt, is dat de 'ernst' van de fout evenredig is met het kwadraat van de fout, dus sneller dan evenredig stijgt naarmate de voorspellingsfout zelf groter is. In hoeverre dit realistisch is hangt af van wat met die voorspellingen in concreto gedaan wordt; daarop kunnen wij hier niet ingaan. Wel merken we op, dat gelijke fouten van verschillende

richting door dit criterium gelijk worden behandeld: of de voorspellingsfout 3% is of —3%, in beide gevallen geven zij een bijdrage van 9 tot de kwadratensom.

Men kan dan per sector de gemiddelde kwadratische voorspellingsfout berekenen voor alle voorspellingen één jaar vooruit, alle voorspellingen twee jaren vooruit, enz., tot aan tien jaren vooruit toe. Doen wij dit voor sector 13, dan is het resultaat als volgt:

Aantal jaren vooruit:

1	2	3	4	5	6	7	8	9	10
19	38	58	55	50	60	69	108	118	131

Duidelijk blijkt, dat de gemiddelde kwadratische voorspellingsfout oploopt, dus dat de voorspellingen volgens het kwadratisch criterium slechter worden, naarmate we verder vooruit voorspellen.

Aldus zijn we erin geslaagd de voorspellingskwaliteit voor een enkele sector met 10 getallen te beschrijven. Zo kunnen we doorgaan met de 26 andere sectoren, waardoor we dus de voorspellingskwaliteit van het materiaal als geheel met 27 × 10 = 270 getallen beschrijven; hetgeen t.o.v. het oorspronkelijke materiaal (1485 voorspellingsfouten) een aanzienlijke reductie is, maar toch nog altijd een stevig aantal. Het ligt daarom voor de hand, dat we proberen nog verder te comprimeren. Daarbij moet men bovendien bedenken, dat een aantal van onze gemiddelde kwadratische voorspellingsfouten op maar heel weinig waarnemingen gebaseerd zijn. Neem bijv. de voorspellingen 10 jaar vooruit. We hebben gezien, dat dit er per sector steeds één is, nl. die voor 1958 verkregen met behulp van de 1948-tabel. Passen we de regel ‘alle fouten kwadrateren, optellen en delen door hun aantal’ in dat geval toe, dan komt dit er eenvoudig op neer, dat we het kwadraat van die ene voorspellingsfout berekenen! Dit is het extreme geval. Van de voorspellingen 9 jaar vooruit hebben we per sector er steeds twee, nl. die voor 1957 gebaseerd op de 1948-tabel en die voor 1958 gebaseerd op de 1949-tabel. Het is duidelijk, dat onze uitspraken inzake de kwaliteit van de voorspellingen voor een groot aantal jaren vooruit gebaseerd zijn op een gering aantal waarnemingen, dus nogal onsolide zijn, zodat het aanbeveling verdient daar iets aan te doen.

Dat is niet bijzonder moeilijk wanneer onze belangstelling uitgaat naar de voorspellingskwaliteit van het materiaal als geheel, dus niet van de afzonderlijke sectoren. We kunnen dan immers middelen over de 27 sectoren! Neem als voorbeeld de voorspellingen één jaar vooruit. Per sector hebben we voor al die voorspellingen de gemiddelde kwadratische voorspellingsfout berekend. Berekent men vervolgens het gemiddelde van al die 27 waarden,[1] dan krijgen we de gemiddelde kwadratische voorspellingsfout van alle voorspellingen één jaar vooruit van alle 27 sectoren tezamen. Dit kan men herhalen voor twee jaren vooruit, enz., en het resultaat is de volgende rij van uitkomsten voor de gemiddelde kwadratische voorspellingsfout voor alle 27 sectoren tezamen:

Aantal jaren vooruit:

1	2	3	4	5	6	7	8	9	10
62	131	189	241	294	361	399	473	560	611

De stijging is bijzonder regelmatig: voor ieder extra jaar vooruit loopt de gemiddelde kwadratische voorspellingsfout met ongeveer 60 op. Vergelijken wij deze uitkomsten met de overeenkomstige van sector 13 (zie blz. 151), dan valt het op, dat deze sector lagere waarden heeft en dus beter voorspeld wordt dan het gemiddelde van alle 27 sectoren. Er is inderdaad sprake van duidelijke verschillen tussen de sectoren onderling, maar het zou te ver voeren daar in detail op in te gaan. Volstaan zij met de constatering, dat over het algemeen de sectoren met grote leveringen relatief beter voorspeld worden dan de minder belangrijke sectoren. De steenkolenmijnen maken daar een uitzondering op. Dit vindt zijn oorzaak in de aanzienlijke verschuiving van kolen naar andere brandstoffen.

We hebben hiermee de mate van nauwkeurigheid van de aan- en afvoervoorspellingen kwantitatief omschreven en bovendien aangegeven hoe die afhankelijk is van de periode waar-

1. De lezer zal zich wellicht afvragen: waarom het gemiddelde, terwijl eerst tegen de gemiddelde voorspellingsfout zoveel bezwaar is gemaakt? Het antwoord luidt, dat het belangrijkste bezwaar tegen de gemiddelde voorspellingsfout de compensatie van positieve en negatieve waarden is – een bezwaar dat voor de gemiddelde kwadratische voorspellingsfouten niet van toepassing is, omdat die positief zijn.

over vooruit voorspeld wordt. Het ligt in de rede dat men zich afvraagt, of *dezelfde* nauwkeurigheid ook bereikt zou kunnen worden met eenvoudiger methoden. Het maken van een aan- en afvoertabel is niet een zaak van een handomdraai en de daaraan verbonden kosten zijn van voorspellingsoogpunt uit bezien alleen verantwoord, wanneer de voorspellingen ook merkbaar beter zijn dan die van meer elementaire, goedkopere methoden. Het gaat hier dus om een relatieve waardering: over de vraag hoe groot de voorspellingsfouten zijn in vergelijking met die van eenvoudiger methoden. In feite hebben wij ons hier al mee bezig gehouden, toen we de resultaten van de investeringsenquête vergeleken met de simpele regel 'geen verandering'. Hier zullen we op een ietwat meer verfijnde manier te werk gaan. Beschouw dan een willekeurige sector en laat in 1948 zijn totale productie voor 60% uit onderlinge leveringen bestaan. Dus zijn die leveringen in dat jaar $60 / 40 = 1^1/_2$ maal zo groot als het bedrag, waarvoor de consumenten de producten van die sector kopen. De 'naïeve' voorspellingsmethode die wij zullen bezien gaat er dan van uit, dat deze verhouding ook op latere jaren van toepassing is. D.w.z., de onderlinge leveringen van die sector in 1949 en latere jaren wordt dan voorspeld als $1^1/_2$ maal de consumptieve bestedingen van die jaren bij die sector. Dit is natuurlijk meer verfijnd dan de regel 'de onderlinge leveringen veranderen niet in de tijd'; integendeel, zij worden voorspeld evenredig met de consumptieve bestedingen op en neer te gaan. Aan de andere kant is de procedure veel minder verfijnd dan de aan- en afvoermethode, omdat geen rekening wordt gehouden met het stromenschema van de diverse sectoren. En hij is veel goedkoper: een volledige tabel is niet vereist, het is voldoende te weten hoeveel procent van de totale productie per sector naar de consument gaat.

Is die goedkopere methode inderdaad veel slechter in het voorspellen? Deze vraag kan men beantwoorden door met deze methode op precies dezelfde manier te werk te gaan, dus door ook gemiddelde kwadratische voorspellingsfouten te berekenen, die dan vervolgens worden vergeleken met die van de aan- en afvoermethode. Het blijkt dan, dat de laatste methode het wint zolang de tabellen niet te zeer verouderd zijn. Om een voorbeeld te nemen: de 1948-tabel is duidelijk beter (althans voor alle sectoren gemiddeld genomen) dan de verhoudingscijfers van 1948 waarmee de naïeve methode werkt. Maar wan-

neer het bijv. zou gaan om de voorspelling van de onderlinge leveringen in 1953 en wanneer de 1948-tabel de meest recente is die ter beschikking staat, dan is het nog niet gezegd dat er geen verhoudingscijfers zijn voor na 1948. Het is denkbaar dat de naïeve methode gebaseerd kan worden op de verhoudingscijfers van 1950, misschien 1951. Dat betekent dat de naïeve methode, hoewel naïever dan de aan- en afvoermethode, toch het voordeel heeft met recenter gegevens te werken. Wanneer dit verschil in beschikbaarheid, dus de tijdsafstand tussen de meest recente tabel en de meest recente verhoudingscijfers, 2 à 3 jaar gaat bedragen, dan verdwijnt daarmee de superioriteit van de aan- en afvoermethode. Dit is een belangwekkend resultaat, dat als een richtlijn kan worden beschouwd voor het tempo waarin de tabellen moeten worden gemaakt.

LITERATUUR

Over de aan wetenschappelijke voorspellingen te stellen eisen heeft wijlen de Amsterdamse hoogleraar Van Dantzig gesproken op de Statistische Dag van 1952, georganiseerd door de Vereniging voor Statistiek [1]. Zie ook Hoofdstuk 2 van Theil [2]. De investeringsenquête van het C.B.S. wordt jaarlijks gepubliceerd [3]; meer details over het onderzoek naar de voorspellingskwaliteit van deze enquête vindt men bij Mouchart e.a. [4]. Zie voorts [5, 6, 7, 8] voor de onderzoekingen van Lips en Schouten, Van de Panne en Rey en Tilanus inzake de voorspellingswaarde van econometrische macro-modellen en van aan- en afvoertabellen. In [8] wordt het probleem van de prijzen in de aan- en afvoeranalyse behandeld. De klassieke theorie gaat er nl. van uit, dat productie, onderlinge leveringen, enz. uitgedrukt zijn in guldens van constante koopkracht. Dit is niet het geval bij de C.B.S.-tabellen (deze zijn in 'lopende' prijzen berekend) en in de genoemde studie wordt gepoogd hierin m.b.v. prijsindices te voorzien. Dit blijkt de voorspellingskwaliteit niet te verbeteren. Zie Christ [9] voor de in de tekst genoemde minder succesvolle Amerikaanse modelvoorspellingen.

Bij de behandeling van het verschijnsel van onderschatte veranderingen zijn de prognoses van sportjournalisten ter sprake gekomen. Hierover kan men lezen in Jochems [10], die deze voorspellingen in zijn dissertatie [11] heeft vergeleken met qua structuur analoge voorspellingen in de zgn. conjunctuurtest (een schriftelijke enquête onder ondernemers betreffende plannen, verwachtingen en realisaties inzake productie, afzet, prijzen, enz., die in de laatste jaren in vele Europese landen opgang heeft gemaakt).

[1] Dantzig, D. van, 'Voorspelling en profetie'. *Statistica Neerlandica*, Jaargang 6 (1952), pp. 195-203.

[2] Theil, H., *Economic Forecasts and Policy*. Second Edition. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam. 1961.

[3] Centraal Bureau voor de Statistiek, *Investeringen in vaste activa.* Uitgeversmaatschappij W. de Haan, Zeist. (Dit is een jaarlijkse publicatie.)

[4] Mouchart, M., H. Theil and J. I. Vorst, 'On the Predictive Value of Investment Surveys'. *Statistica Neerlandica*, Jaargang 17 (1963), pp. 287-297.

[5] Lips, J., and D. B. J. Schouten, 'The Reliability of the Policy Model Used by the Central Planning Bureau of the Netherlands'. Hoofdstuk 2 van *Income and Wealth*, Series VI, edited bij M. Gilbert and R. Stone. Bowes & Bowes, London. 1957.

[6] Panne, C. van de, 'De voorspellingskwaliteit van de Centrale Economische Plannen, 1949-1956'. *De Economist*, Jaargang 107 (1959), pp. 91-123.

[7] Rey, G., and C. B. Tilanus, 'Input-Output Forecasts for The Netherlands, 1949-1958'. *Econometrica*, Vol. 31 (1963), pp. 454-463.

[8] Tilanus, C. B., and G. Rey, 'Input-Output Volume and Value Predictions for The Netherlands, 1948-1958'. Rapport 6309 van het Econometrisch Instituut van de Nederlandsche Economische Hoogeschool (1963).

[9] Christ, C. F., 'A Test of an Econometric Model for the United States, 1921-1947'. Pagina's 35-107 van *Conference on Business Cycles*. National Bureau of Economic Research, New York. 1951.

[10] Jochems, D. B., 'Forecasting the Outcomes of Soccer Matches: A Statistical Appraisal of the Dutch Experience'. *Metrika*, Vol. 5 (1962), pp. 194-207.

[11] Jochems, D. B., *Economische weerberichten. Enige empirische onderzoekingen met behulp van conjunctuurtestgegevens.* Dissertatie Rotterdam. 1962.

De 14 endogene variabelen, waarvan voorwaardelijke voorspellingen in dit hoofdstuk werden bezien, zijn de volgende: invoer van goederen, consumptie, bruto investeringen, nationaal product, werkgelegenheid, indirecte belastingen, loon- en inkomstenbelasting betaald door loontrekkers, inkomstenbelasting betaald door nietloontrekkers, het inkomen van die groep en prijsindices voor de volgende vijf bestedingscategorieën: export van goederen, consumptie, investeringen, voorraden en overheidsuitgaven aan goederen. Zie [5] voor verdere details.

De 23 variabelen, waarvan de onvoorwaardelijke voorspellingen werden bezien, zijn: waarde en volume van de consumptie, waarde en volume van in- en uitvoer van goederen, de waarde van de in- en

uitvoer van diensten, het saldo van de dienstverleningsbalans, netto-investeringen, toegevoegde waarde in bedrijven, overheidslonen en -salarissen, overheidsuitgaven aan goederen, indirecte belastingen, industriële productie, productie van de bouwnijverheid, beroepsbevolking, arbeidsproductiviteit, werkgelegenheid in bedrijven en bij de overheid en de prijsindices van consumptie en van in- en uitvoer van goederen. Zie [6] voor verdere details.

6. OVER ONZEKERHEID EN WAARSCHIJNLIJKHEID

1. ONZEKERHEID: EEN BIJNA UNIVERSEEL VERSCHIJNSEL

In de voorgaande hoofdstukken waren er tal van passages, waar we deden alsof een bepaalde uitspraak met zekerheid kan worden gegeven, terwijl dit in feite toch maar in zeer beperkte mate het geval is. Bijvoorbeeld, bij de behandeling van lineaire programmering maakten wij op blz. 40 kennis met een ondernemer, wiens afzet per kwartaal van tevoren gegeven is. In feite pleegt men zijn toekomstige afzet niet met volmaakte nauwkeurigheid te kennen. Een ander voorbeeld is dat van het kritieke pad. Op blz. 70 is gesteld dat men de tijd benodigd voor de afzonderlijke bewerkingen dient te bepalen. Maar kàn men die met volmaakte nauwkeurigheid bepalen? Een derde voorbeeld is dat van de econometrische modellen van Hoofdstuk 4. Er zijn verscheidene manieren om het mechanisme van een volkshuishouding met een model af te beelden. Welke is de meest adequate? Dat is een probleem van econometrisch onderzoek, waarbij rekening dient te worden gehouden met de aard van de volkshuishouding, met het doel waarvoor het model wordt geconstrueerd, met het ter beschikking staande statistische materiaal, enz. Een aprioristische uitspraak is hier niet gemakkelijk te geven. En we hebben in Hoofdstuk 5 gezien, dat hoe het model ook luidt, er geen kwestie van is dat het de realiteit op volmaakte wijze zal weergeven. Zelfs het meest voortreffelijke model schiet tekort als ons criterium neerkomt op een foutloze weergave van de werkelijkheid.

Nu is dit alles op zichzelf niets bijzonders. Ook in de natuurwetenschappen worden we geconfronteerd met afwijkingen tussen model en realiteit. Nemen we bijv. een ijzeren staaf en verwarmen we die gelijkmatig, dan luidt het model dat iedere verdere graad temperatuurverhoging leidt tot dezelfde uitzetting. Maar meten we het na, dan vinden we dat het model niet precies klopt. Dit zal gedeeltelijk aan waarnemingsfouten liggen, gedeeltelijk aan het feit dat we er niet in geslaagd zijn de staaf volledig gelijkmatig te verwarmen, gedeeltelijk ook aan meer fundamentele oorzaken.

De omstandigheid, dat de zoveel oudere natuurwetenschappen met dezelfde kwaal van afwijkingen tussen de wiskundig geformuleerde theorie enerzijds en het waarnemingsmateriaal anderzijds te maken hebben, doet de vraag rijzen of we iets van hun ervaring kunnen leren. Welnu, leren kunnen we altijd, maar een belangrijk punt is dat de afwijkingen op economisch terrein (en dus ook de onzekerheid op dat terrein) omvangrijker plegen te zijn. Bovendien is men in de natuurwetenschappen meestal in staat door herhaald experimenteren tot nieuwe uitkomsten van hetzelfde verschijnsel te komen; en neemt men het gemiddelde van al deze uitkomsten, dan mag men althans in de meeste gevallen aannemen dat zo'n gemiddelde een betere benadering is dan een enkele uitkomst. In de economie is het herhaalde experiment maar op beperkte schaal mogelijk. Men kan natuurlijk wel iets doen wat hierop lijkt. Keren we bijv. terug naar de ondernemer van blz. 40, wiens afzet per kwartaal gegeven werd verondersteld. Is aan die veronderstelling niet voldaan, dan kan men bijv. aldus te werk gaan: men bepaalt voor een aantal jaren de verhouding van de afzet van het tweede kwartaal tot die van het eerste kwartaal, berekent het gemiddelde van al die verhoudingen; en vervolgens past men dit gemiddelde toe om aan het einde van het eerste kwartaal in het huidige jaar een voorspelling te maken omtrent de afzet in het tweede kwartaal van dit jaar. Dit kan in een aantal gevallen een heel redelijke methode zijn; maar men dient zich geen illusies te maken omtrent de omvang van de afwijkingen die van deze methode het gevolg zullen zijn. En past de ondernemer de op blz. 46 beschreven methode toe op basis van dergelijke afzetvoorspellingen, dan is het lang niet zeker of die methode zijn optimale karakter zal behouden.

2. DE GEDACHTENGANG VAN DE WAARSCHIJNLIJKHEIDSREKENING

Gegeven dus de belangrijke rol van de onzekerheid op de terreinen waarmee wij ons hier bezig houden, is het zaak dat wij dit probleem niet omzeilen maar op een serieuze wijze aanpakken. Dit nu is op verschillende manieren mogelijk. Op blz. 36 werden we geconfronteerd met een situatie, waar de onzekerheid inderdaad aanzienlijk pleegt te zijn, nl. wedden op Duin-

digt. We pakten het probleem aldus aan: ons geld werd verdeeld over de diverse paarden op een dusdanige manier dat de kleinste winst nog zo groot mogelijk is. Anders gezegd, we weten niet welk paard zal winnen – als we het zouden weten, zouden we natuurlijk al ons geld op dat ene paard zetten –; we verkeren dus in het onzekere en gaan daarom ons geld over de diverse kanshebbers verdelen volgens het bovengenoemde criterium.

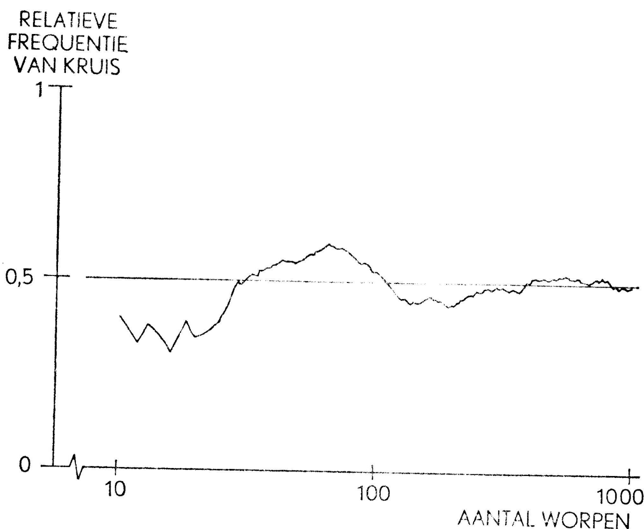
Het woord 'kanshebbers' werd hier met opzet geïntroduceerd. Men kan zich nl. afvragen of het gehanteerde criterium wel gebruik maakt van onze kennis van zaken betreffende paardenraces. (Laten we hier gemakshalve aannemen, dat een dergelijke kennis van zaken ons niet ontzegd kan worden.) Stel bijv., dat we menen er redelijk vertrouwen in te kunnen hebben, dat Bruin het er beter af zal brengen dan Schimmel. Ons criterium maakt op geen enkele wijze gebruik van deze kennis van zaken en men kan zich afvragen in hoeverre dat verstandig is. Wij zullen hierop niet verder ingaan maar ons concentreren op het volgende aspect: blijkens ons voorbeeld is het althans in zekere gevallen mogelijk een uitspraak te doen van het type 'Bruin heeft een grotere kans om te winnen dan Schimmel'. Dit soort uitspraken wordt in het dagelijks leven op grote schaal gedaan, al hebben de termen 'kans' en 'waarschijnlijkheid' dan natuurlijk niet de precieze betekenis die zij in de waarschijnlijkheidsrekening hebben. Om misverstand te voorkomen willen wij de principiële tegenstanders van de gokzucht erop attent maken, dat de waarschijnlijkheidsrekening er niet in de eerste plaats voor is om een wetenschappelijk fundament te leveren voor hetgeen zij zozeer verfoeien. Wellicht is de schijn tegen ons dank zij Duindigt; ook is het waar, dat de oorsprong van de waarschijnlijkheidsrekening gezocht moet worden bij de adviezen, die Franse edellieden in de zeventiende eeuw aan prominente wiskundigen vroegen teneinde de opbrengsten van hun spelen te verbeteren. Het is ook nog waar, dat wij in de volgende pagina's af en toe een dobbelsteen zullen laten rollen; maar dat representeert alleen maar een poging om een zo eenvoudig mogelijke illustratie te geven, zulks op precies dezelfde manier als waarop wij appels en pannekoeken op de lagere school in stukken hebben gesneden om het beginsel van de breuken te illustreren en, later op de middelbare school, hebben gerekend met vorderingen en schulden teneinde het beginsel van de nega-

tieve getallen te doorgronden. De toepassingen van de waarschijnlijkheidsrekening zijn groot in aantal en de volgende opsomming (ontleend aan Feller's bekende *Introduction to Probability Theory and Its Applications*) moge dit illustreren: Een munt eenmaal opgooien, een munt 100 maal opgooien, drie munten tegelijkertijd opgooien, een stel kaarten schudden, roulette spelen, de levensduur van een radioactief atoom of van een persoon bepalen, een steekproef van personen nemen en het aantal linkshandigen daarvan bepalen; en ook het geslacht van een jonggeborene, het aantal bezette lijnen in een telefooncentrale, het aantal keren dat een bepaald nummer wordt opgebeld, ruis in een electrisch communicatiestelsel, de kwaliteitsbeheersing van een productieproces, de frequentie van ongelukken, het aantal dubbelsterren in een bepaalde hemelstreek, de plaats van een deeltje dat aan diffusie onderhevig is. En dat is nog maar een greep uit het totaal van alle toepassingen, hetgeen de lezer in klimmende mate zal inzien naarmate hij de Epiloog van dit boek nadert.

De beginselen van de klassieke waarschijnlijkheidsrekening kunnen het eenvoudigst worden uitgelegd aan de hand van een simpel experiment. Stel we hebben een zuivere, d.w.z. volmaakt symmetrisch geconstrueerde munt; stel voorts dat we die 10 maal opgooien. Dan zal bijv. het resultaat zijn: 4 maal kruis, 6 maal munt, hetgeen we plegen uit te drukken door te zeggen dat de *relatieve frequentie* van kruis 4 op 10 is, dus 0,4. Stel vervolgens, dat we niet 10 maar 100 keer opgooien. Laat het resultaat dan zijn: 54 maal kruis, 46 maal munt, zodat in dit geval de relatieve frequentie van kruis gelijk is aan 0,54. Stel tenslotte, dat we de moed hebben om 10.000 keer te gooien met bijv. het volgende resultaat: 5060 maal kruis en 4940 maal munt, dus de relatieve frequentie van kruis is no 0,506.

Wij willen niet beweren deze experimenten zelf uitgevoerd te hebben, maar wel dat dit soort numerieke resultaten ruwweg correspondeert met hetgeen verwacht mag worden wanneer men ze inderdaad zou uitvoeren (en wanneer de munt ook werkelijk zuiver is). Waar het om gaat is dat de relatieve frequentie van kruis meer en meer nadert tot de waarde $^1/_2$ naarmate men vaker en vaker dezelfde munt opgooit. Begrijp goed, het 'pad' naar $^1/_2$ is bepaald niet geleidelijk of regelmatig. Het is best denkbaar, dat we na de eerste reeks van 10 worpen (met uitslag: 4 maal kruis, 6 maal munt) een tweede reeks hebben

met als resultaat: 3 maal kruis, 7 maal munt. Voor de eerste 20 worpen hebben we dan 4 + 3 = 7 maal kruis en als relatieve frequentie 7 / 20 = 0,35, dus nog verder van $^1/_2$ dan de 0,4 van de eerste 10 worpen. Maar uiteindelijk, hoe onregelmatig het pad ook is, het is vrijwel zeker dat we bij $^1/_2$ terecht zullen komen. De volgende figuur geeft een typisch beeld van een dergelijk pad:[1]



Deze onvolmaakte regelmaat is de basis van de klassieke waarschijnlijkheidsrekening. Onvolmaakt wat betreft de ups en downs van het pad; maar regelmaat wat betreft het feit, dat uiteindelijk de waarde $^1/_2$ meer en meer benaderd wordt. Dit laatste drukt men uit door te zeggen, dat de 'waarschijnlijkheid van kruis' gelijk is aan $^1/_2$. Een waarschijnlijkheid (ook wel kans genoemd) is steeds een getal tussen 0 en 1, de grenzen daarbij inbegrepen. Is een gebeurtenis zeker, dan is zijn waarschijnlijkheid per definitie gelijk 1; is het zeker dat hij nièt zal optreden, dan is zijn waarschijnlijkheid 0. Een voorbeeld van een zekere gebeurtenis is 'hetzij kruis hetzij munt' (aannemende dat de munt niet op zijn rand kan blijven staan).

1. Teneinde de figuur niet al te onhandzaam te doen zijn, is langs de horizontale as een logarithmische schaalverdeling aangebracht.

De waarschijnlijkheidsrekening (of kansrekening) is in feite niets anders dan landmeting op het terrein van de onzekerheid. Is een gebeurtenis zeker, dan kent de waarschijnlijkheidstheoreticus haar de kans 1 toe; is zij bijna zeker, dan is de toegekende kans bijna 1, bijv. 0,98; is het bijna zeker dat de gebeurtenis niet zal optreden, dan zal de toegekende kans een klein positief getal zijn, bijv. 0,05. Enzovoorts. Ten aanzien van de wijze, waarop de waarschijnlijkheidstheoreticus kanswaarden aan de diverse gebeurtenissen toekent, bestaat een zekere mate van vrijheid, maar daaraan zijn uiteraard beperkingen verbonden. Zo behoeft hij er niet noodzakelijkerwijs van uit te gaan, dat een munt zuiver is. Het is denkbaar, dat de kans op kruis 0,52 is en dit kan blijken bij een voldoende groot aantal worpen. Maar wél zal men er in het algemeen van uitgaan, dat die kans constant is in de loop van de tijd, dus niet varieert van worp tot worp. Kort gezegd, het komt erop neer dat men enerzijds de werkelijkheid geen geweld wenst aan te doen door waarschijnlijkheden te postuleren die te ver van de waargenomen relatieve frequenties afwijken, anderzijds zal men de waarschijnlijkheidsspecificatie mathematisch niet zo ingewikkeld willen maken dat er vrijwel niet mee te werken is. Daarnaast kent de kansrekening ook zijn eigen regels, geheel los dus van het terrein van toepassing. Dit zijn de zgn. *axioma's*, te vergelijken met 'door twee punten gaat één en niet meer dan één rechte lijn' in de meetkunde. Dit zijn er drie:

(1) Een kans is steeds een getal tussen 0 en 1 (zie boven).

(2) De kans op een *zekere* gebeurtenis, die dus gegarandeerd optreedt, is 1 (zie boven).

(3) Als twee gebeurtenissen niet tegelijkertijd kunnen optreden, is de kans dat één van de twee optreedt gelijk aan de som van de twee afzonderlijke kansen.

Dit laatste axioma behoeft wellicht enige verduidelijking. Stel daartoe, dat we een symmetrisch geconstrueerde dobbelsteen hebben. Werpen we hem op, dan zijn er zes mogelijkheden: 1 oog, 2 ogen, . . ., 6 ogen. Stel nu, dat we hem 60 keer opwerpen met bijv. het volgende resultaat:

1 oog	2 ogen	3 ogen	4 ogen	5 ogen	6 ogen
8	10	11	9	13	9

De relatieve frequentie van '1 oog' is dus 8 / 60 = 0,133; van

'2 ogen' is hij 10 / 60 = 0,167; en van '1 of 2 ogen' is hij (8 + 10) : 60 = 0,3, dus de som van de afzonderlinge relatieve frequenties van '1 oog' en '2 ogen'. Werpen we niet 60 maar 6000 keer, dan vinden we bijv. het volgende:

1 oog	2 ogen	3 ogen	4 ogen	5 ogen	6 ogen
1020	1010	964	985	965	1056

De relatieve frequentie van '1 oog' is nu 1020 / 6000 = 0,170 en van '2 ogen' 1010 / 6000 = 0,168. Aldus zien we, evenals in het geval van kruis of munt, dat de relatieve frequentie van deze twee gebeurtenissen 'naar een limiet convergeert', in casu in beide gevallen naar de waarschijnlijkheid 1 / 6. Maar we zien ook, dat de relatieve frequentie van '1 oog of 2 ogen' nu gelijk is aan

(1020 + 1010) : 6000 = 0,338;

hij is wederom gelijk aan de som van de relatieve frequenties van '1 oog' en '2 ogen' en convergeert naar de waarde $^1/_3$, d.w.z. naar de som van de waarschijnlijkheden van '1 oog' en '2 ogen'. En dat is nu juist wat het derde axioma op het oog heeft. De gebeurtenissen '1 oog' en '2 ogen' kunnen niet tegelijkertijd optreden wanneer men een dobbelsteen opgooit; het resultaat is immers hetzij 1 oog hetzij 2 ogen hetzij ... hetzij 6 ogen, nooit bijv. 3 en 4 ogen tegelijkertijd. Het axioma zegt, dat in een dergelijk geval een combinatie van twee gebeurtenissen een kans heeft, die gevonden wordt door de kansen van de afzonderlijke gebeurtenissen op te tellen. Welnu, dit blijkt geheel in overeenstemming te zijn met hetgeen geldt voor de relatieve frequenties, die we aan de waarschijnlijkheden ten grondslag hebben gelegd.[1]

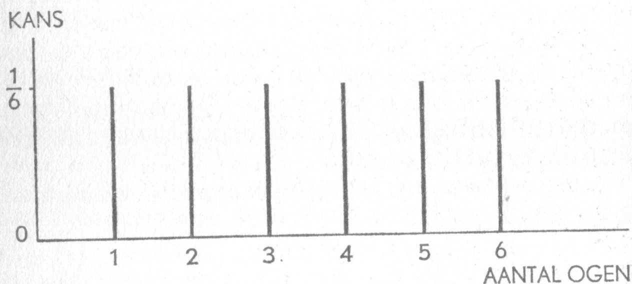
De drie axioma's vormen het wiskundige fundament van het

1. Kunnen twee gebeurtenissen *A* en *B* wèl tegelijkertijd optreden, dan is de kans op 'hetzij *A* hetzij *B*' kleiner dan de som van de kansen op *A* en *B* afzonderlijk. Neem daartoe het dobbelsteenvoorbeeld en laat *A* zijn de gebeurtenis 'hetzij 1 oog hetzij 2 ogen' en *B* 'hetzij 2 hetzij 3 ogen'. Dan kunnen *A* en *B* tegelijkertijd optreden, nl. wanneer 2 ogen worden geworpen. Voorts is 'hetzij *A* hetzij *B*' gelijkwaardig met 'hetzij 1 hetzij 2 hetzij 3 ogen' en de kans daarop is (vanwege het derde axioma!) $^1/_6 + ^1/_6 + ^1/_6 = ^1/_2$. Maar de kans op *A* is $^1/_3$ en die op *B* ook; hun som is $^2/_3$, dus groter dan de kans op 'hetzij *A* hetzij *B*'.

omvangrijke bouwwerk, dat waarschijnlijkheidsrekening heet. We zullen nu de steigers beklimmen, maar niet verder dan noodzakelijk is voor de doeleinden van dit boek.

3. DISCRETE VERDELINGEN

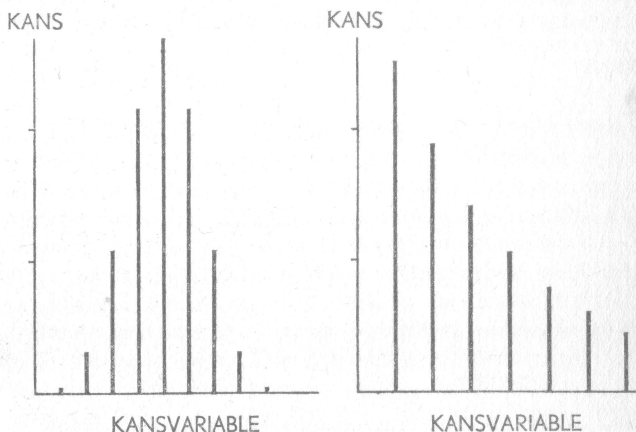
Blijkens het citaat uit Feller's boek houdt de waarschijnlijkheidsrekening zich met een aanzienlijke variëteit van verschijnselen bezig. Sommige daarvan zijn kwalitatief van aard, d.w.z. ze worden niet met behulp van een getal gemeten. Een voorbeeld is het geslacht van een jonggeborene, dus mannelijk of vrouwelijk. (Voor belangstellenden vermelden we, dat de kans op een jongen in Nederland ruim 0,51 is.) Wij zullen ons echter bepalen tot numerieke kansverschijnselen, die dus wèl met behulp van een getal worden gemeten. Een voorbeeld is de dobbelsteen, wanneer we uitsluitend letten op het aantal ogen. Dit voorbeeld heeft bovendien tot eigenschap, dat er slechts een beperkt aantal getalswaarden kan worden gerealiseerd (in casu een zestal). In dat geval spreken we van een 'discrete verdeling', hetgeen het type is waartoe we ons in eerste instantie zullen beperken. Voor ons dobbelsteenvoorbeeld kunnen we deze verdeling op eenvoudige wijze grafisch weergeven:



Horizontaal noteren we de waarden van het verschijnsel waar het om gaat, in dit geval het aantal ogen; verticaal meten we de kans waarmee deze waarden in feite worden gerealiseerd. Aldus ontstaat het beeld van zes verticale staven, elk ter lengte van $^1/_6$.

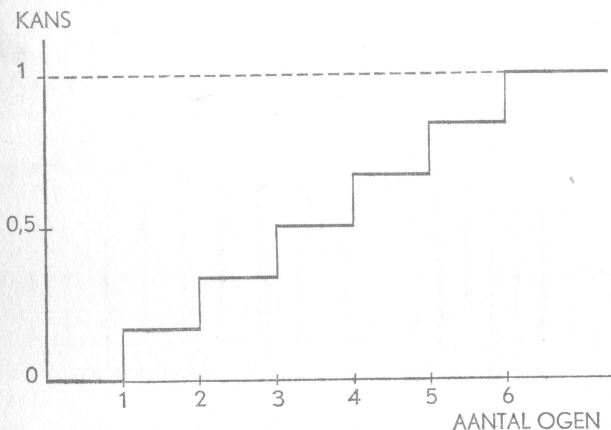
Het dobbelsteenvoorbeeld is natuurlijk maar een van de vele

denkbare toepassingen; de betekenis van de waarschijnlijkheidsrekening (evenals trouwens die van andere onderdelen van de wiskunde) is juist gelegen in het feit, dat er toepassingen op zo veel verschillende terreinen zijn. Onder deze omstandigheden is het noodzakelijk te werken met een terminologie, die los staat van de concrete toepassing. Wij zullen daarom de horizontaal afgezette variabele verder met *kansvariabele* aanduiden; hier dus het aantal ogen, maar het kan ook zijn het aantal geboorten in Nederland per jaar of de afzet van een of ander product of het aantal afgekeurde radio's dat per dag een lopende band verlaat, enz. Het is overigens bepaald niet zo, dat de waarschijnlijkheden van de door een kansvariabele aangenomen waarden alle dezelfde moeten zijn, zoals voor ons dobbelsteenvoorbeeld geldt (alle $^1/_6$). Dat is niet meer dan een bijzonder geval en in het algemeen zullen de staven dan ook niet alle even lang zijn, zoals in de volgende figuren wordt geïllustreerd:



De hier besproken figuren geven weer wat bekend staat als de *massafunctie* van de kansvariabele. Deze functie specificeert dus de waarschijnlijkheid van elk van de waarden die deze variabele kan aannemen. Er is nog een andere manier om een verdeling te karakteriseren, nl. die van de *verdelingsfunctie*. Daartoe bezien we wederom het dobbelsteenvoorbeeld en we vragen ons nu af: hoe groot is de kans, dat het aantal gegooide ogen ten hoogste 0 is, ten hoogste 1, ten hoogste 2, enz. Welnu, de

kans op een aantal ogen van ten hoogste 0 is gelijk nul, want het is zeker dat er minimaal 1 oog wordt gegooid. De kans op een aantal ogen van ten hoogste 1 is gelijk $^1/_6$; er is immers maar één manier om dit resultaat te bereiken en dat is door 1 oog te gooien. De kans om ten hoogste 2 ogen te gooien is $^1/_3$; dit wordt gerealiseerd door *hetzij* 1 oog *hetzij* 2 ogen te gooien, en aangezien deze twee gebeurtenissen niet gelijktijdig kunnen optreden tellen we volgens het derde axioma hun kansen (elk $^1/_6$) op. Analoog is de kans om ten hoogste 3 ogen te gooien gelijk aan $^1/_2$; dan moeten we nogmaals $^1/_6$ bijtellen (nl. voor de gebeurtenis '3 ogen'). Aldus gaan we voort tot de waarschijnlijkheid 1 wordt bereikt voor een aantal gegooide ogen van ten hoogste 6; een nog groter aantal ogen is er immers niet. Het beeld van een dergelijke verdelingsfunctie is dat van een trapjeslijn, klimmend van links naar rechts. Linksonder start hij bij 0, rechtsboven eindigt hij bij 1, en de sprongen worden telkens gemaakt bij die waarden die de kansvariabele in kwestie met positieve waarschijnlijkheid aanneemt.



4. VERWACHTING EN VARIANTIE

De tot nu toe besproken verdelingen worden volledig gespecificeerd door op te geven: (i) de waarden die de kansvariabele kan aannemen (in ons voorbeeld, 1, 2, . . ., 6 ogen) en (ii) de

bijbehorende waarschijnlijkheden. Een dergelijke karakterisering kan echter omslachtig zijn wanneer de kansvariabele een groot aantal waarden kan aannemen en in het bijzonder wanneer de waarschijnlijkheden niet, zoals in ons voorbeeld, alle gelijk zijn. Aldus ligt het voor de hand te zoeken naar een klein aantal maatstaven, die de belangrijkste eigenschappen van een verdeling beschrijven. Het bekendst is de *mathematische verwachting* of kortweg de verwachting van een verdeling. Dat is niet zo iets als een toekomstverwachting, maar eerder een maatstaf voor hetgeen men gemiddeld genomen als uitkomst dient te verwachten. Neem weer het dobbelsteenvoorbeeld: er zijn zes mogelijke waarden, 1, 2, . . ., 6, alle even waarschijnlijk, dus nemen we gewoon het gemiddelde:

$$(1 + 2 + \ldots + 6) : 6 = 3^1/_2.$$

Algemeen: stel dat we een kansvariabele hebben die de volgende waarden kan aannemen:

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad \ldots$$

en laat de bijbehorende waarschijnlijkheden zijn:

$$p_1 \quad p_2 \quad p_3 \quad p_4 \quad \ldots,$$

dan is de verwachting gedefinieerd als

$$\mu = p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 + \ldots.$$

Links staat μ (een Griekse letter, spreek uit 'mu'), het algemeen gebruikte symbool van een mathematische verwachting. Blijkens de definitie wordt de verwachting als volgt gevonden: neem alle waarden $x_1, x_2, \ldots$ die de kansvariabele kan aannemen, ga ze wegen (vermenigvuldigen) met de bijbehorende kansen (p_1, $p_2, \ldots$) en tel dan alles op. Aldus is μ een 'gewogen gemiddelde' van alle waarden $x_1, x_2, \ldots$. Merk op, dat deze algemene definitie de uitkomst $3^1/_2$ van het dobbelsteenvoorbeeld als een bijzonder geval omvat. Die uitkomst is nl. verkregen door die algemene definitie toe te passen: neem alle aantallen ogen die geworpen kunnen worden, (1, 2, . . ., 6), ga ze wegen met de bijbehorende waarschijnlijkheden (alle $^1/_6$ in dit geval) en tel alles op:

$$\mu = {}^1/_6 \times 1 + {}^1/_6 \times 2 + \ldots + {}^1/_6 \times 6 =$$
$$(1 + 2 + \ldots + 6) : 6 = 3^1/_2.$$

Het is duidelijk, dat de zojuist ingevoerde p's alle groter dan of gelijk aan nul moeten zijn; ze zijn immers waarschijnlijkheden, zie het eerste axioma. Ook geldt, dat de som van alle p's gelijk moet zijn aan 1:

$$p_1 + p_2 + p_3 + p_4 + \ldots = 1,$$

hetgeen als volgt valt in te zien. Het is kennelijk onmogelijk, dat een tweetal waarden (bijv. x_1 en x_2) gelijktijdig worden gerealiseerd. Dan volgt uit het derde axioma, dat de kans op 'hetzij x_1 hetzij x_2' gevonden wordt door optelling; die kans is dus gelijk aan $p_1 + p_2$. Oòk is het waar, dat 'hetzij x_1 hetzij x_2' onmogelijk gelijktijdig met x_3 gerealiseerd kan worden; dus is de kans op 'hetzij x_1 hetzij x_2 hetzij x_3' gelijk aan $p_1 + p_2 + p_3$. Zo kunnen we doorgaan. Op het laatst krijgen we de gebeurtenis

'hetzij x_1 hetzij x_2 hetzij x_3 hetzij x_4 hetzij . . .',

dus de gebeurtenis dat één van alle x-en wordt gerealiseerd, en de kans daarop is gelijk aan de som van alle p's. Maar het is zèker dat één van alle x-waarden wordt gerealiseerd en blijkens het tweede axioma is de kans dan 1; en dus, zoals beweerd, moet de som van alle waarschijnlijkheden (de p's) van een discrete verdeling steeds gelijk zijn aan 1. Het dobbelsteenvoorbeeld voldoet aan deze eis; de som van de p's wordt dan immers eenvoudig gevonden door zes maal $^1/_6$ te nemen.

De verwachting van een verdeling is een voor de hand liggende maatstaf wanneer het erom gaat de 'centrale tendentie' te karakteriseren, dus om aan te geven hetgeen men gemiddeld genomen als uitkomst zal mogen verwachten. Maar in feite kan de werkelijke uitkomst aanzienlijk van de verwachting afwijken. Het dobbelsteenvoorbeeld illustreert dit duidelijk genoeg. De verwachtingswaarde is $3^1/_2$ oog en wanneer we 3 of 4 ogen werpen zitten we daar dicht bij.[1] Maar we kunnen ook wel 1 oog werpen en de kans daarop is $^1/_6$, dus bepaald niet zonder meer te verwaarlozen. Het ligt daarom in de rede, dat men een maatstaf wenst te gebruiken voor de omvang van de afwijkingen van de verwachtingswaarde. De bekendste is de *variantie*, vaak aangeduid met σ^2 (spreek uit 'sigma kwadraat'). Dit is een maatstaf voor spreiding: nul als de kansvariabele met zekerheid gelijk is aan zijn verwachting, positief wanneer de variabele verscheidene waarden kan aannemen, en groter en groter naarmate die waarden verder van elkaar komen te liggen. Preciezer, de variantie is een gewogen gemiddelde – dus net als bij de verwachting – en wel een gewogen gemiddelde

1. In dit geval kàn de kansvariabele zelfs niet met de verwachting samenvallen, want $3^1/_2$ oog kan niet gegooid worden. Dit is vergelijkbaar met de uitspraak 'de Nederlandse gezinnen hebben gemiddeld $2^1/_4$ kinderen'.

van de gekwadrateerde verschillen tussen de aangenomen waarden en de verwachting. In het dobbelsteenvoorbeeld zijn die verschillen:

$1 - 3^1/_2 = -2^1/_2, \quad 2 - 3^1/_2 = -1^1/_2, \quad \ldots,$
$6 - 3^1/_2 = 2^1/_2.$

Gaan we kwadrateren en het gewogen gemiddelde nemen (met de waarschijnlijkheden als gewichten), dan vinden we:

$$\sigma^2 = {}^1/_6(-2^1/_2)^2 + {}^1/_6(-1^1/_2)^2 + \ldots + {}^1/_6(2^1/_2)^2 = 2^{11}/_{12}.$$

Algemener: zijn de waarden die de kansvariabele aanneemt $x_1, x_2, \ldots$ en de bijbehorende waarschijnlijkheden $p_1, p_2, \ldots$, dan is de variantie:

$$\sigma^2 = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \ldots,$$

waarbij uiteraard $\mu = p_1x_1 + p_2x_2 + \ldots$, de verwachting. We zien gemakkelijk, dat een variantie nooit negatief kan zijn, dus altijd positief of nul. Het gaat immers om kwadraten, elk vermenigvuldigd met een bepaalde kans, en die kans kan evenmin negatief zijn. Ook zien we gemakkelijk, dat de variantie nul is zodra $x_1, x_2, \ldots$ alle met μ samenvallen. De kansvariabele neemt dan met zekerheid de waarde μ aan.

De vierkantswortel van de variantie (σ dus) staat bekend als de *standaarddeviatie* van de verdeling, die evenals de variantie zelf een vaak gebruikte maatstaf voor de spreiding der uitkomsten is. Voor het dobbelsteenvoorbeeld neemt hij de waarde 1,71 aan (tot op 2 decimalen nauwkeurig). Resumerend kunnen wij dus stellen, dat bij het werpen van een zuivere dobbelsteen het aantal gegooide ogen verdeeld is rond een verwachting van $3^1/_2$ met een standaarddeviatie van 1,71.

5. HET REKENEN MET VERWACHTINGEN

Het gebruik van de mathematische verwachting voor de beschrijving van de centrale tendentie van een verdeling biedt zekere voordelen, wanneer het gaat om grootheden die we uit onze kansvariabelen kunnen afleiden; en dat komt vaak voor. Nemen we daartoe wederom onze dobbelsteen; laten we veronderstellen, dat de ogen nu dubbel geteld moeten worden. De mogelijke waarden zijn dan:

2 4 6 8 10 12;

immers, wordt 1 oog gegooid, dan telt het resultaat als 2; worden 2 ogen gegooid, dan telt het resultaat als 4, enz. De

waarschijnlijkheden zijn kennelijk ongewijzigd, dus $^1/_6$. Dit betekent, dat de mathematische verwachting nu wordt

$$^1/_6 \times 2 + {}^1/_6 \times 4 + \ldots + {}^1/_6 \times 12 = 7,$$

dus het dubbele van de oorspronkelijke verwachting. Een resultaat van dit type geldt algemeen. Neem daartoe een kansvariabele die de waarden $x_1, x_2, \ldots$ aanneemt met de waarschijnlijkheden, $p_1, p_2, \ldots$; dan is, zoals we hebben gezien, de verwachting gelijk aan $p_1x_1 + p_2x_2 + \ldots$. Neem nu aan, dat de kansvariabele vermenigvuldigd wordt met een constante factor k; dan neemt hij de waarden, $kx_1, kx_2, \ldots$ aan met precies dezelfde waarschijnlijkheden $p_1, p_2, \ldots$. De nieuwe verwachting wordt

$$p_1(kx_1) + p_2(kx_2) + \ldots = k(p_1x_1 + p_2x_2 + \ldots),$$

dus k maal de oorspronkelijke verwachting.

Stel vervolgens, dat we er een gewoonte van maken om bij het werpen van een dobbelsteen aan het aantal gegooide ogen een tiental toe te voegen. De mogelijke uitkomsten worden dan

11 12 13 14 15 16,

elk met een kans $^1/_6$. De mathematische verwachting is nu:

$$^1/_6 \times 11 + {}^1/_6 \times 12 + \ldots + {}^1/_6 \times 16 = 13^1/_2,$$

dus 10 hoger dan de oorspronkelijke verwachting. Ook dit geldt algemeen. Neemt de kansvariabele de waarden, x_1, x_2, ... aan met waarschijnlijkheden $p_1, p_2, \ldots$, en verhogen wij hem met c (zodat de waarden $x_1 + c$, $x_2 + c$, ... worden), dan vinden we voor de nieuwe verwachting:

$$p_1(x_1 + c) + p_2(x_2 + c) + \ldots = p_1x_1 + p_2x_2 + \ldots$$
$$+ c(p_1 + p_2 + \ldots) = p_1x_1 + p_2x_2 + \ldots + c,$$

dus de oorspronkelijke verwachting verhoogd met c. Merk op, dat wij hierbij gebruik hebben gemaakt van de eigenschap dat de som van alle p's gelijk is aan 1.

Het is mogelijk deze resultaten nog verder uit te breiden. Neem een tweetal kansvariabelen, bijv. de aantallen ogen gegooid met twee dobbelstenen of de aantallen millimeters regen in de eerste en de tweede helft van het jaar of de verkopen van een bedrijf in twee opeenvolgende jaren. We kunnen dergelijke kansvariabelen optellen en krijgen dan een nieuwe kansvariabele: het totale aantal ogen met twee dobbelstenen gegooid, de regenval in het hele jaar, de totale verkoop van een twee-jaarsperiode. We kunnen van deze nieuwe kansvariabele de verwachting berekenen; we kunnen hetzelfde doen met de twee oorspronkelijke kansvariabelen. Het blijkt dan, dat de verwachting van de som gelijk is aan de som van de twee afzonder-

lijke verwachtingen. Dus is de verwachting van het totaal aantal ogen gegooid met twee dobbelstenen gelijk aan de som van de verwachtingen voor de afzonderlijke dobbelstenen (dus $3^1/_2 + 3^1/_2 = 7$ als de dobbelstenen zuiver zijn). En analoog voor de regenval en de verkoop. De afleiding wordt in het Aanhangsel van dit hoofdstuk gegeven. Het resultaat kan nog verder uitgebreid worden: wanneer we een willekeurig aantal kansvariabelen hebben (niet noodzakelijk twee) en die alle optellen, dan vinden we de verwachting van de som door de verwachtingen van alle afzonderlijke componenten op te tellen. En tenslotte: neem een willekeurige lineaire combinatie van kansvariabelen, dus k_1 maal de eerste variabele plus k_2 maal de tweede plus k_3 maal de derde, enz., waarbij k_1, k_2, k_3, . . . alle constante vermenigvuldigingsfactoren zijn (bijv. 6, —3, 1, . . .). Het gaat om de verwachting van die lineaire combinatie. Nu hebben we hierboven gezien, dat als we een kansvariabele met een constante factor k_1 vermenigvuldigen, het resultaat is een nieuwe kansvariabele met een verwachting k_1 maal zo groot. Anders gezegd, geven we de verwachting van de eerste kansvariabele aan met μ_1, dan verkrijgen we na vermenigvuldiging een nieuwe variabele met verwachting $k_1\mu_1$. Analoog heeft de tweede kansvariabele na vermenigvuldiging met k_2 een verwachting $k_2\mu_2$, de derde $k_3\mu_3$, enz. In de lineaire combinatie worden al deze vermenigvuldigde kansvariabelen bij elkaar opgeteld; en we vermeldden zojuist, dat de verwachting van een som gelijk is aan de som der afzonderlijke verwachtingen. Dus is de verwachting van de lineaire combinatie gelijk aan

$$k_1\mu_1 + k_2\mu_2 + k_3\mu_3 + \dots .$$

In woorden: de verwachting van een willekeurige lineaire combinatie van kansvariabelen is gelijk aan dezelfde lineaire combinatie (met dezelfde gewichten k_1, k_2, . . .!) van de verwachtingen van de afzonderlijke kansvariabelen.

Dit is een resultaat, dat de 'algebra der verwachtingen' bijzonder vergemakkelijkt. Het wordt dan ook in de waarschijnlijkheidsrekening op grote schaal gebruikt.

6. OVER CONTINUE VERDELINGEN

We hebben ons tot nu toe bepaald tot kansvariabelen die slechts bepaalde waarden (meestal gehele waarden: 0, 1, 2,

. . .) kunnen aannemen. Daar zijn talrijke voorbeelden van; niet alleen het aantal ogen van een dobbelsteen maar ook het aantal ongevallen per week op een bepaalde weg, het aantal slachtoffers per jaar van een bepaalde ziekte, het aantal auto's van een rijschool dat per maand defect raakt, het aantal auto's dat op vrijdag om 17.30 uur in de Rotterdamse tunneltraverse in de rij staat te wachten, enz. Maar er zijn ook gevallen waarvoor dit niet opgaat. Preciezer, er zijn kansvariabelen die in beginsel alle denkbare waarden kunnen aannemen, althans alle waarden van een zeker interval (bijv. alle positieve waarden). Voorbeelden: de lengte van een dienstplichtige op het moment dat hij voor de militaire dienst wordt gekeurd, het gewicht van de man op hetzelfde tijdstip, zijn gewichtstoename gedurende de tijd van zijn militaire dienst, de tijd die een auto in de tunneltraverse moet doorbrengen wanneer hij op vrijdagmiddag tussen 17.15 en 17.30 daarin verzeild raakt. Uiteraard kunnen we deze gevallen terugbrengen tot het discrete type door voldoende af te ronden. Wanneer we bijv. de lengte van dienstplichtigen in gehele aantallen decimeters meten, dan hebben we een kansvariabele die behoudens verwaarloosde uitzonderingen de waarden 15 (corresponderend met een lengte tussen 145 en 155 centimeter), 16, 17, . . ., 22 aanneemt. Maar door aldus te werk te gaan doen we het kansverschijnsel eigenlijk geweld aan en het verdient daarom aanbeveling het probleem bij de wortel aan te pakken.

Het gaat hier om de zgn. 'continu' verdeelde kansvariabelen in tegenstelling tot die van de discrete verdelingen. Het volgende voorbeeld is instructief. Stel we hebben op willekeurige wijze 50 Nederlandse gezinnen geselecteerd voor een budgetonderzoek. Alle bestaan uit man, vrouw en twee kinderen; alle hebben een maandinkomen van f 1000. Het gaat om de bestedingen van deze gezinnen; in hetgeen volgt zullen wij ons uitsluitend interesseren voor de totale consumptieve uitgaven per gezin. Het spreekt vanzelf, dat ook al is de gezinssamenstelling dezelfde en ook al hebben alle gezinnen hetzelfde inkomen, de uitgaven niet hetzelfde zullen zijn en dus een zekere spreiding zullen vertonen. Smaken verschillen nu eenmaal; het ene gezin is wat conservatief en spaart op grote schaal, het andere laat de genoegens van het heden zwaar wegen en spaart niet of nauwelijks; het ene gezin heeft in het onderzochte kwartaal toevallig zware lasten (doktersrekeningen, versleten autokoppeling),

voor het andere gezin geldt dit niet, enz. Kortom, wij beschouwen de consumptieve uitgaven van een Nederlands gezin bestaande uit man, vrouw en twee kinderen met een inkomen van f 1000 per maand als een kansvariabele. Laten we de 50 gezinnen, die ons gegevens hebben verstrekt, nummeren naar de alfabetische volgorde van hun achternaam; laten de 50 uitgaafbedragen (in guldens per maand) als volgt zijn:

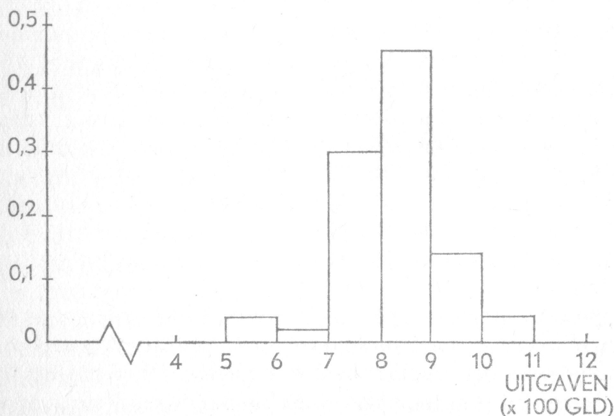
1.	716	14.	797	27.	768	40.	827
2.	937	15.	759	28.	818	41.	939
3.	782	16.	862	29.	808	42.	786
4.	835	17.	883	30.	998	43.	724
5.	1082	18.	722	31.	891	44.	837
6.	1012	19.	803	32.	768	45.	572
7.	701	20.	570	33.	623	46.	836
8.	834	21.	878	34.	816	47.	937
9.	864	22.	866	35.	879	48.	904
10.	811	23.	766	36.	820	49.	733
11.	798	24.	916	37.	931	50.	839
12.	714	25.	876	38.	835		
13.	801	26.	700	39.	874		

Uitkomsten van dit type zijn in beginsel niet anders dan die, welke we verkrijgen door 50 maal een dobbelsteen op te gooien en elke keer het resultaat neer te schrijven. Voor ons huidige probleem is het belangrijkste verschil, dat de resultaten van de dobbelsteen slechts zes alternatieve waarden kunnen zijn (1, 2, . . ., 6 ogen), terwijl de maandelijkse consumptieve uitgaven in beginsel elk willekeurig bedrag kunnen aannemen.[1]

Het komt de overzichtelijkheid ten goede wanneer we de resultaten grafisch weergeven. Wij constateren, dat er geen gezinnen zijn, die minder dan f 500 uitgaven; dat er twee zijn, die

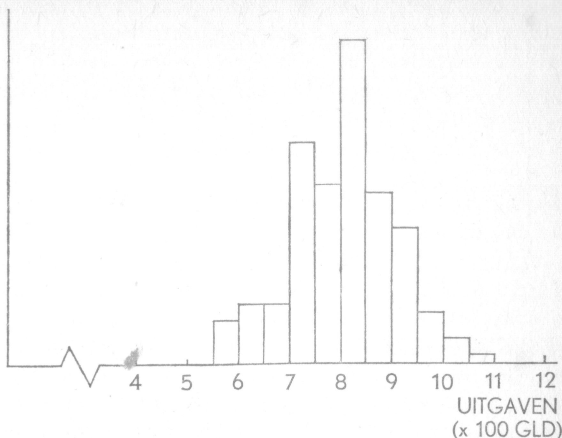
1. De spitsvondige lezer zal opmerken, dat de uitgaven steeds een veelvoud van een cent moeten zijn en dus toch discreet verdeeld zijn! Dit is correct, maar het is de gewoonte om het discrete karakter van een kansvariabele te verwaarlozen als de sprongetjes zo gering zijn vergeleken met de variabiliteit van het kansverschijnsel. Dit is hier inderdaad het geval: een cent is weinig vergeleken met de verschillen in uitgaven, die hier soms een paar honderd gulden zijn. (Wij hebben de uitgaven in gehele guldens uitgedrukt om het beeld eenvoudig te houden.)

tussen f 500 en f 600 uitgaven (No. 20 en 45); dat er één is, die tussen f 600 en f 700 uitgaf (No. 33). Enzovoorts. De relatieve frequentie van uitgaven beneden f 500 is dus nul, van tussen f 500 en f 600 is 2 / 50 = 0,04 en van tussen f 600 en f 700 is 1 / 50 = 0,02, enz. We maken nu een diagram met horizontaal de uitgaven en verticaal de relatieve frequentie waarmee de gezinnen in kwestie een uitgaafbedrag hebben, dat correspondeert met de langs de horizontale as aangegeven intervallen. Het resultaat is als volgt:



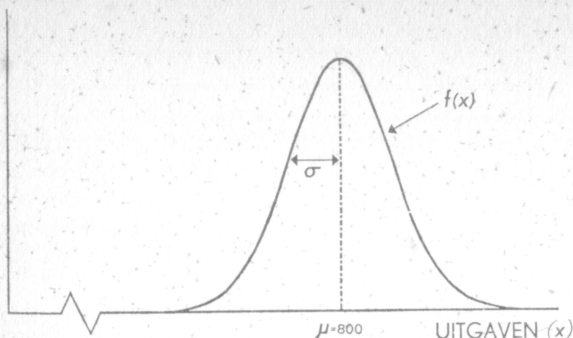
Men ziet, dat de grootste frequentie ligt rond een bedrag van ruim f 800 en dat de frequentie naar links en naar rechts, gemeten vanaf dit centrale bedrag, min of meer geleidelijk daalt.

Stel nu, dat we niet 50 maar 150 gezinnen hebben, alle bestaande uit man, vrouw en twee kinderen en alle met een maandinkomen van f 1000. Dan kunnen we precies dezelfde procedure op deze grotere groep toepassen. Dit leidt tot een lijst van 150 uitgaafbedragen (hier niet gereproduceerd) en het resultaat kan wederom in een diagram worden afgebeeld. Het grotere aantal waarnemingen stelt ons in staat met een fijnere klasse-indeling te werken; het beeld dat hierdoor ontstaat treft men aan op blz. 175, bovenaan. In grote lijnen is het beeld hetzelfde als dat van het vorige diagram, maar er is nu een klein beetje meer regelmaat.



Het is nu van groot belang in te zien, dat wat wij hier op het ogenblik doen in feite niets anders is dan wat wij in het begin van dit hoofdstuk deden toen wij in voortdurend grotere aantallen onze munt opwierpen. Daar bestudeerden we de relatieve frequentie van kruis en het resultaat was: 0,4 bij 10 keer opgooien; 0,54 bij 100 keer; 0,506 bij 10000 keer; en wij constateerden dat die relatieve frequentie convergeert naar een vaste waarde, genaamd de waarschijnlijkheid van kruis. Hier ligt de zaak geheel analoog. Wij bestuderen de verdeling van de consumptieve uitgaven van een bepaald type van gezinnen (man, vrouw en twee kinderen, inkomen f 1000 per maand) en dat doen wij door een voortdurend groter aantal gezinnen aan een budgetonderzoek te onderwerpen. Ook hier vinden wij een stijgende regelmaat, een convergeren naar een limiet, die voor het onderhavige geval is afgebeeld op de volgende bladzijde. (Een zeer bekende figuur in de waarschijnlijkheidsrekening.) De getekende curve is die van een *dichtheidsfunctie* (ook wel verdelingsdichtheid of kortweg dichtheid genoemd) en wel de dichtheidsfunctie van een zgn. *normale verdeling*. De vergelijking van de dichtheid van deze verdeling luidt:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-{}^1/_2(x-\mu)^2/\sigma^2}$$



een nogal ingewikkelde formule, waarin 5 symbolen voorkomen: x, π, e, μ en σ. Hiervan staat x voor de verschillende numerieke waarden die de kansvariabele (in ons voorbeeld: de uitgaven) kan aannemen. Verder zijn π en e gewoon bepaalde getallen: $\pi = 3{,}14159\ldots$ (de bekende verhouding van de omtrek van een cirkel tot zijn middellijn) en $e = 2{,}71828\ldots$[1] De symbolen μ en σ suggereren dat deze de verwachting resp. de standaarddeviatie weergeven; dit vermoeden is inderdaad juist. Men leidt de verwachting en de variantie van een continue verdeling af door een procedure, die in de wiskunde bekend staat als 'integreren', maar hierop zullen wij niet nader ingaan. De regels voor het rekenen met verwachtingen, die wij hierboven voor discrete verdelingen hebben afgeleid, zijn ook op continue verdelingen van toepassing.

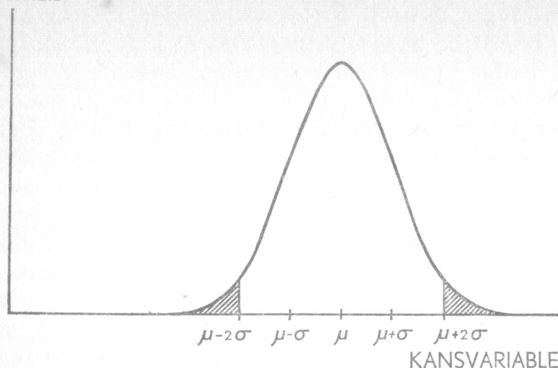
De gedaante van de dichtheid van een normale verdeling is die van een klokvormige curve met een top in het midden en een symmetrisch verloop naar beide zijden. Wanneer we de curve aflopen van links naar rechts, dan vinden wij dat hij dicht bij de as begint en stijgt en wel voortdurend sneller stijgt, althans in het begin. Dan komen we op een punt waarop hij nog wel degelijk stijgt maar ophoudt sneller te stijgen. Dat punt wordt op de horizontale as gevonden door van de verwachting

1. Het getal e staat bekend als het grondtal van het zgn. natuurlijke logarithmenstelsel en kan als volgt als een limiet gedefinieerd worden. Neem enig positief geheel getal n; bepaal de n de macht van $1 + 1/n$ en neem dan de limiet voor n naar oneindig. Voor $n = 1$ vinden we 2; voor $n = 2$ is het resultaat 2,25; voor $n = 3$ wordt het 2,37...; enz.

(μ) de standaarddeviatie (σ) af te trekken. De curve blijft ook daarna stijgen, zij het in een vertraagd tempo, tot hij zijn maximum bereikt bij de verwachting. Daarna gaat hij dalen, langzaam in het begin maar voortdurend sneller. Het grootste daaltempo wordt bereikt op het punt, dat correspondeert met de verwachting plus de standaarddeviatie. Daarna daalt hij voortdurend langzamer en vlijt zich tegen de horizontale as aan. De curve blijft altijd boven de as. De waarde nul bereikt de dichtheid pas 'in het oneindige' (plus oneindig en min oneindig), d.w.z. nooit. De lezer zal dan wellicht tegenwerpen, dat het toch maar een kwalijke zaak is, dat wij het uitgaafpatroon van gezinshuishoudingen hebben beschreven met een curve die negatieve uitgaven tot aan min oneindig toelaat. Wij hebben alle respect voor zijn spitsvondigheid en moeten toegeven, dat een negatief bedrag voor de totale consumptieve uitgaven van een gezin niet goed denkbaar is; maar we moeten op onze beurt tegenwerpen, dat het bezwaar niet veel practische betekenis heeft. De hier gebruikte normale verdeling heeft een verwachting van 800 gulden per maand en een standaarddeviatie van 100 gulden per maand. Met behulp van een tabel van de normale verdeling vindt men, dat de kans op een negatieve waarde dan ongeveer 6×10^{-16} is (nul komma en dan 15 nullen en dan een 6). Anders gezegd, wanneer we een ploeg onderzoekers zouden hebben die per dag een miljoen (10^6) gezinnen op hun uitgaven onderzoeken en die dat een miljard (10^9) dagen zouden volhouden, dus bijna 30.000 eeuwen, dan zou er volgens het model wel eens een keer een gezin tussen kunnen zitten met negatieve uitgaven. Wanneer onze modellen niet vaker tot moeilijkheden aanleiding geven mogen we tevreden zijn. We willen eerlijk toegeven, dat we wèl vaker moeilijkheden hebben (die echter in de meeste gevallen van andere aard zijn).

Zojuist spraken we over een tabel van de normale verdeling en inderdaad zijn er van deze verdeling uitvoerige tabellen gemaakt, die men in meer of minder uitvoerige vorm aantreft in boeken over waarschijnlijkheidsrekening of statistiek. Wij gaan hier niet in detail op in maar bepalen ons tot één aspect, dat aan de hand van de figuur op blz. 178 wordt geïllustreerd.

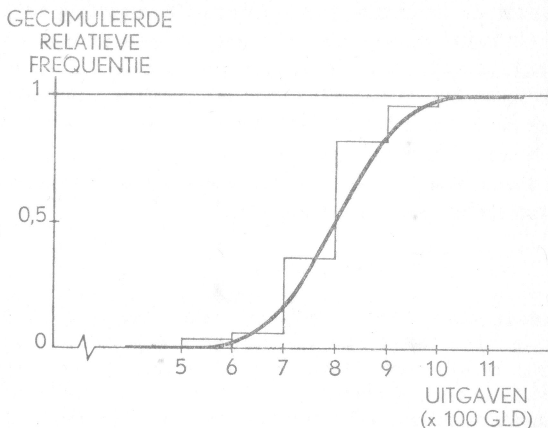
Het gaat hierom: wat is de kans, dat we een zeker veelvoud van de standaarddeviatie, bijv. tweemaal de standaarddeviatie, of



nóg verder van de verwachting terecht komen? Kennelijk zitten we dan op de horizontale as hetzij rechts van $\mu + 2\sigma$ hetzij links van $\mu — 2\sigma$. Het antwoord is al in de figuur aangegeven: die waarschijnlijkheid is gelijk aan de som van de twee gearceerde oppervlakten. In een getal: die kans is gelijk aan 0,0455 (dus iets beneden $^1/_{20}$) en dat geldt voor *alle* normale verdelingen, wat hun verwachting en standaarddeviatie ook mogen zijn. In gewoon Nederlands: is een kansvariabele normaal verdeeld, dan komt het iets minder vaak dan 1 op 20 voor, dat de afwijking van de verwachtingswaarde gelijk is aan tweemaal de standaarddeviatie of nog groter. Vooral in statistische toepassingen wordt van deze eigenschap op grote schaal gebruik gemaakt.

De dichtheid van een continu verdeelde kansvariabele speelt in grote trekken dezelfde rol als de massafunctie van een discreet verdeelde variabele. Nu hebben we indertijd gezien, dat de verdelingsfunctie uit de massafunctie verkregen werd door 'cumulatie van links'; we vroegen immers successievelijk naar de kans op ten hoogste 0, 1, 2, . . . ogen. Een analoge vraagstelling is mogelijk voor een continue verdeling. Laten we beginnen met de 50 gezinnen wier uitgaafbedragen op blz. 173 zijn opgesomd. Het laagste bedrag is f 570 (corresponderend met No. 20) en dus is de relatieve frequentie $^1/_{50}$ voor een uitgaafbedrag van ten hoogste 570 gulden. Het een-na-laagste bedrag is f 572 (No. 45), dus stijgt de relatieve frequentie plotseling tot $^1/_{25}$ voor een uitgaafbedrag van ten hoogste 572 gulden. Aldus

kunnen wij voortgaan tot de relatieve frequentie 1 wordt bij een uitgaafbedrag van ten hoogste 1082 gulden; dit is immers het maximum dat door de groep van 50 is uitgegeven (nl. door No. 5). Het resultaat wordt als in het discrete geval met een trapjeslijn weergegeven:



Door de trapjeslijn heen vindt men een vloeiende stijgende kromme. Dit is de limiet waartoe de trapjeslijn nadert, wanneer men niet 50 maar een onbepaald groot aantal gezinnen zou nemen. Die limietcurve correspondeert uiteraard met de al eerder getekende dichtheid van de normale verdeling en wordt, analoog aan het discrete geval, de verdelingsfunctie genoemd. Het zal duidelijk zijn, dat de verdelingsfunctie van een continue verdeling in het algemeen voortdurend stijgt en bepaald niet de allure van een trapjeslijn heeft, eerder die van een glijbaan.

In het bovenstaande hebben we de normale verdeling zeer centraal geplaatst. Dat is tot op zekere hoogte gerechtvaardigd. Niet alleen komt de normale verdeling (of althans iets dat er veel op lijkt) bij feitelijk waarnemingsmateriaal tamelijk vaak voor, maar bovendien speelt hij een grote rol in de wiskundig-statistische theorie. De belangrijkste reden van dit laatste is de zgn. Centrale Limietstelling, die zegt dat als we een aantal kansvariabelen optellen, de som bij benadering normaal verdeeld is als dat aantal maar groot genoeg is. Aan deze stelling zijn zekere voorwaarden verbonden; de belangrijkste is wel

deze, dat de afzonderlijke kansvariabelen niet op enigerlei wijze systematisch mogen samenhangen.[1] Aan deze eis is bijv. niet voldaan wanneer het gaat om de lengte en het gewicht van dezelfde persoon; want wie lang is weegt ook meestal tamelijk zwaar. (Dat er ook lange mensen met gering gewicht zijn betekent nog niet dat er geen samenhang is!)

Toch wordt de betekenis van de normale verdeling soms overschat. Er is wel eens gezegd, dat wiskundigen denken dat het empirisch vaststaat, dat de dingen van de wereld normaal verdeeld zijn, en dat de beoefenaars van de toegepaste statistiek denken dat het een wiskundige wet is dat het zo moet zijn. Het behoeft niet zo te zijn. Bij de behandeling van de wachttijdtheorie in Hoofdstuk 9 zullen we met niet-normaal verdeelde continue variabelen geconfronteerd worden.

LITERATUUR

Er is een groot aantal leerboeken op het gebied van de waarschijnlijkheidsrekening, waarvan vele zich ook met statistische toepassingen bezig houden. Feller's boek [1] is al in de tekst genoemd; het is bijzonder goed, maar niet werkelijk elementair. Het boek van de Utrechtse hoogleraar Freudenthal [2] is iets eenvoudiger. De boeken van Moroney [3] en Wijvekate [4] richten zich i.h.b. op statistische toepassingen en zijn beide elementair.

[1] Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. I. Second Edition. John Wiley and Sons, Inc., New York. 1957.

[2] Freudenthal, H., *Waarschijnlijkheidsrekening en statistiek.* Tweede druk. Volksuniversiteitsbibliotheek (Tweede reeks, No. 57), De Erven F. Bohn, Haarlem. 1962.

[3] Moroney, M. J., *Facts from Figures.* Third Revised Edition. Pelican Book A 236, Penguin Books, Harmondsworth, Middlesex. 1956.

[4] Wijvekate, M. L., *Verklarende statistiek.* Aulaboek Nr. 39, Het Spectrum, Utrecht. 1960.

AANHANGSEL

Stel we hebben twee kansvariabelen, weer te geven door X_1 en X_2. Bijvoorbeeld, X_1 kan zijn het aantal ogen gegooid met één dobbel-

1. Preciezer: de kansvariabelen moeten 'onafhankelijk verdeeld' zijn. Hierop wordt in het Aanhangsel van dit hoofdstuk nader ingegaan.

steen en X_2 het aantal gegooid met een tweede dobbelsteen; of X_1 is het inkomen van een gezin en X_2 het bedrag aan voedingsuitgaven van hetzelfde gezin. Wij zullen ons bepalen tot discrete verdelingen, maar de resultaten van dit Aanhangsel zijn evenzeer van toepassing op continue verdelingen.

We nemen een eenvoudig voorbeeld, waarin X_1 drie waarden kan aannemen, nl. 1, 2 en 3, en X_2 twee: 1 en 2. Dan zijn er zes mogelijkheden in totaal: we kunnen hebben $X_1 = 1$, $X_2 = 1$, weer te geven door (1, 1); verder $X_1 = 1$, $X_2 = 2$, weer te geven door (1, 2); ook $X_1 = 2$, $X_2 = 1$, weer te geven door (2, 1), enz. Alle zes mogelijkheden tezamen kunnen op eenvoudige wijze in het volgende schema gerangschikt worden:

	$X_2 = 1$	$X_2 = 2$
$X_1 = 1$	(1, 1)	(1, 2)
$X_1 = 2$	(2, 1)	(2, 2)
$X_1 = 3$	(3, 1)	(3, 2)

Hierbij stelt het eerste getal tussen haken dus steeds de uitkomst van de eerste kansvariabele (X_1) voor en het tweede getal de uitkomst van de tweede kansvariabele (X_2). Om het wat concreet voor te stellen: we werpen twee dobbelstenen en de uitkomst van de eerste bepaalt X_1 volgens regel:

1 of 2 ogen: $X_1 = 1$
3 ogen: $X_1 = 2$
4, 5 of 6 ogen: $X_1 = 3$;

en de uitkomst van de tweede dobbelsteen bepaalt X_2:

1 oog: $X_2 = 1$
2 of meer ogen: $X_2 = 2$.

In dit soort situaties wordt het van belang de waarschijnlijkheden van alle *paren* van uitkomsten te specificeren. D.w.z., we dienen ons bezig te houden met de kans op (1, 1), op (1, 2), enz. Dat zullen wij nu doen op een willekeurige manier, daarbij de zojuist gegeven dobbelsteeninterpretatie volledig vergetend. Laat dan de kans op (1, 1) gelijk zijn aan 0,2. Dat betekent dus, dat de waarschijnlijkheid 0,2 is dat de eerste kansvariabele gelijk is aan 1 en bovendien tegelijkertijd de tweede variabele ook gelijk aan 1. Laat de kans op (1, 2) gelijk zijn aan 0,1, enz. Het gemakkelijkst noteert men die waarschijnlijkheden in hetzelfde schema als hetwelk we zojuist gebruikten, dat we nu echter ter onderscheid volledig met lijnen (dus ook rechts en onder) afsluiten:

	$X_2 = 1$	$X_2 = 2$
$X_1 = 1$	0,2	0,1
$X_1 = 2$	0,3	0,1
$X_1 = 3$	0,1	0,2

We merkten op dat die waarschijnlijkheden op willekeurige wijze zouden worden gespecificeerd, maar het spreekt vanzelf dat aan die willekeur grenzen zijn gesteld. In de eerste plaats dienen de kansen alle minimaal 0 en maximaal 1 te zijn (het eerste axioma). In de tweede plaats kunnen de zes combinaties nooit met twee stuks tegelijkertijd gerealiseerd worden en is het zeker dat één van hen wèl gerealiseerd zal worden; en dus (het tweede en derde axioma) moet de som van alle zes waarschijnlijkheden gelijk zijn aan 1.

We zullen ons nu bezig houden met de vraag: hoe groot is de waarschijnlijkheid dat $X_1 = 1$? Deze is niet gegeven maar wel uit de gegevens af te leiden. $X_1 = 1$ kan zich immers maar op twee manieren voordoen, nl. als (1, 1) of als (1, 2) gerealiseerd wordt. Deze uitkomsten kunnen niet gelijktijdig gerealiseerd worden ($X_2 = 1$ en $X_2 = 2$ sluiten elkaar immers uit), dus wordt volgens het derde axioma de kans op $X_1 = 1$ gevonden door optelling van de kansen op (1, 1) resp. (1, 2), dus $0{,}2 + 0{,}1 = 0{,}3$. Analoog vinden we, dat de kansen op $X_1 = 2$ en $X_1 = 3$ gelijk zijn aan 0,4 resp. 0,3. Voor X_2 gaan we op precies dezelfde manier te werk. De kans op $X_2 = 1$ vinden we door optelling van de kansen van (1, 1), (2, 1) en (3, 1), dus $0{,}2 + 0{,}3 + 0{,}1 = 0{,}6$. Het volgende schema is instructief:

	$X_2 = 1$	$X_2 = 2$	Marginaal
$X_1 = 1$	0,2	0,1	0,3
$X_1 = 2$	0,3	0,1	0,4
$X_1 = 3$	0,1	0,2	0,3
Marginaal	0,6	0,4	1

Door horizontale optelling van de waarschijnlijkheden binnen de lijnen vinden we 0,3 en 0,4 en 0,3, zie de meest rechtse kolom; en dat zijn juist de waarschijnlijkheden van $X_1 = 1$, $X_1 = 2$ resp. $X_1 = 3$ zoals we die zojuist hebben afgeleid. Door verticale optelling vinden we 0,6 en 0,4: de waarschijnlijkheden van $X_2 = 1$ resp. $X_2 = 2$. Dit zijn wat men noemt de waarschijnlijkheden van de 'marginale verdeling' van X_1 resp. X_2.[1] Die verdelingen zijn relevant wanneer

1. Marginaal betekent hier: in de marge (rechts in het schema voor X_1, onderaan voor X_2). Dit moet niet verward worden met het marginale begrip in de economie (marginale opbrengst, marginale kosten), dat wiskundig bezien op een differentiaalquotiënt neerkomt.

men elke kansvariabele afzonderlijk beziet, dus niet het tweetal in onderling verband. De 1 rechtsonder in het schema is de som van de onderste rij en ook de som van de meest rechtse kolom. Uiteraard moeten die sommen van waarschijnlijkheden gelijk 1 zijn; het zijn immers de waarschijnlijkheden van twee verdelingen.

De verwachtingen van X_1 resp. X_2 worden nu gedefinieerd als de verwachtingen van hun marginale verdelingen. Voor X_1 wordt dit

$0{,}3 \times 1 + 0{,}4 \times 2 + 0{,}3 \times 3 = 2$

en voor X_2:

$0{,}6 \times 1 + 0{,}4 \times 2 = 1{,}4.$

We gaan nu kijken naar de som, $X_1 + X_2$. Wanneer beide kansvariabelen 1 zijn, dus in het geval (1, 1), dan is die som 2; in de gevallen (1, 2) en (2, 1) is de som 3, enz. We houden weer hetzelfde 3×2 schema aan voor de waarden aangenomen door de som:

	$X_2 = 1$	$X_2 = 2$
$X_1 = 1$	2	3
$X_1 = 2$	3	4
$X_1 = 3$	4	5

Voor elke (X_1, X_2)-combinatie is de waarschijnlijkheid gegeven; we vinden dan de verwachting van de som gemakkelijk:

$0{,}2 \times 2 + 0{,}1 \times 3 + 0{,}3 \times 3 + 0{,}1 \times 4 + 0{,}1 \times 4 + 0{,}2 \times 5 = 3{,}4;$

dus precies de som van de verwachtingen van X_1 en X_2 $(2 + 1{,}4)$.

Het algemene bewijs is niet moeilijk voor wie met somtekens kan werken. Laat X_1 de waarden $a_1, a_2, \ldots$ aannemen en X_2 de waarden $b_1, b_2, \ldots$; schrijf p_{ij} voor de kans op $X_1 = a_i$, $X_2 = b_j$. Dan worden de marginale waarschijnlijkheden van X_1 gevonden door p_{ij} over j te sommeren en die van X_2 door over i te sommeren. De verwachtingen van X_1 en X_2 zijn

$$\sum_i \sum_j p_{ij} a_i \quad \text{resp.} \quad \sum_i \sum_j p_{ij} b_j$$

en de verwachting van $X_1 + X_2$:

$$\sum_i \sum_j p_{ij}(a_i + b_j) = \sum_i \sum_j p_{ij} a_i + \sum_i \sum_j p_{ij} b_j,$$

dus de som van de verwachtingen van X_1 en X_2. De uitbreiding tot drie en meer kansvariabelen is evenmin moeilijk; de p's krijgen meer indices, dat is alles.

Laten we nu de numerieke waarden van onze waarschijnlijkheden iets veranderen:

	$X_2 = 1$	$X_2 = 2$	Marginaal
$X_1 = 1$	0,15	0,1	0,25
$X_1 = 2$	0,15	0,1	0,25
$X_1 = 3$	0,3	0,2	0,5
Marginaal	0,6	0,4	1

Hier is iets bijzonders aan de hand. Zoals tevoren komt $X_2 = 1$ in 60% van de gevallen voor (kans 0,6), maar nu wordt in een kwart hiervan $X_1 = 1$ gerealiseerd, ook in een kwart $X_1 = 2$ en in de resterende helft $X_1 = 3$. Het bijzondere is dat die onderverdeling voor $X_2 = 2$ precies dezelfde is. $X_2 = 2$ komt voor in 40% van de gevallen; gegeven $X_2 = 2$ doet $X_1 = 1$ zich voor in een kwart van de gevallen, idem voor $X_1 = 2$ en $X_1 = 3$ in de helft. Het belangrijke van deze situatie is, dat een eventuele kennis van de waarde die X_2 aanneemt ons nu van geen enkel nut is om een nadere uitspraak omtrent X_1 te doen. In het oorspronkelijke voorbeeld was dit wel het geval. Ook daar is het waar, dat $X_2 = 1$ zich voordoet in 60% van de gevallen; maar als we dit eenmaal weten, dan weten we ook dat $X_1 = 2$ zich in de helft van de gevallen realiseert. Voor $X_2 = 2$ is dit maar een kwart, een belangrijk verschil dus.

Is in de nieuwe situatie kennis van X_2 nutteloos om iets naders over X_1 te weten te komen, omgekeerd is het ook waar, dat kennis omtrent X_1 ons niet helpt om iets naders over X_2 te zeggen. Stel bijv. dat we weten dat $X_1 = 3$; dan zien we dat in 60% van de gevallen $X_2 = 1$ zich zal realiseren, in 40% van de gevallen $X_2 = 2$. Precies dezelfde onderverdeling hebben we als het gegeven is dat $X_1 = 1$ of dat $X_1 = 2$, zodat deze kennis ons inderdaad niets helpt t.o.v. X_2.

In een dergelijke situatie plegen we te zeggen, dat X_1 en X_2 *onafhankelijk verdeeld* zijn. Of dit het geval is kan men nagaan op de hier beschreven wijze, maar de volgende controle is nog iets handiger: in het geval van onafhankelijkheid is iedere waarschijnlijkheid binnen de lijnen gelijk aan het product van de marginale waarschijnlijkheden in de bijbehorende rij en kolom. Voor linksboven vinden we: $0{,}15 = 0{,}25 \times 0{,}6$; voor rechtsboven: $0{,}1 = 0{,}25 \times 0{,}4$; en zo nog vier keer. Het is een zaak van een beetje algebra om aan te tonen dat beide methoden gelijkwaardig zijn.

7. HET STRATEGIEBEGRIP

1. STRATEGIE EN TACTIEK

In het dagelijks spraakgebruik wordt wel gezegd: hij mag een bekwaam tacticus zijn, een groot strateeg is hij niet. Waarmee dan bedoeld wordt, dat de persoon in kwestie de problemen van de dag op handige wijze weet aan te pakken, maar dat hij onvoldoende rekening houdt met de grote lijn en met de consequenties van zijn handelen op lange termijn. Het strategiebegrip dat in dit hoofdstuk aan de orde zal komen heeft zonder twijfel met dit onderscheid te maken, maar het heeft een nauwkeuriger gedefinieerde inhoud.

Het schaakspel biedt een geschikte mogelijkheid om dit begrip te introduceren. Neem daartoe een willekeurige opstelling; laat Wit aan zet zijn. Hij zal een aantal alternatieve zetten overwegen, hij zal zich bij elk van deze alternatieven afvragen wat het antwoord van Zwart zou kunnen zijn, hij zal vervolgens zijn mogelijkheden bij de volgende zet in het oog nemen, enz. Doet hij dit consequent een aantal zetten vooruit, dan maakt hij wat in het schaakspel heet een 'combinatie'.

Een strategie nu is een combinatie in deze zin, maar dan consequent tot het bittere einde. Om dit te illustreren gaan we ervan uit, dat we aan het begin van het spel staan. Wit moet dus zijn eerste zet doen. Hij heeft de keuze uit 20 mogelijkheden; er zijn immers acht pionnen, die elk hetzij een hetzij twee velden kunnen worden opgeschoven, en er zijn twee paarden die elk twee kanten uit kunnen. Dat is alles; de rest zit vast. Wanneer Wit zit te denken over de relatieve merites van deze 20 alternatieven dient hij te denken aan Zwart's repliek. Welnu, ook Zwart heeft bij zijn eerste zet 20 mogelijkheden, nl. precies dezelfde maar dan aan de andere kant van het bord. In totaal zijn er dus $20 \times 20 = 400$ mogelijkheden voor de eerste zet, die van Wit en van Zwart bij elkaar genomen. Vervolgens zal Wit (en bedenk wel, hij moet al deze mogelijkheden overwegen vóór de eerste zet te doen!) zich moeten bezinnen op zijn repliek bij de tweede zet. Dan zijn er nog veel meer mogelijkheden: elke pion die nog niet verzet is kan een of twee plaatsen naar voren; als er een pion bij de eerste zet verplaatst is kan die een plaats naar voren (tenzij hij pal tegenover een zwarte pion staat); paarden

kunnen springen; een loper of toren of koningin of koning kan naar voren indien bij de eerste zet vrijgemaakt door pionverschuiving. Het totaal aantal mogelijkheden van eerste zet Wit plus eerste zet Zwart plus tweede zet Wit zal dan omstreeks 10.000 belopen. En dat is nog maar het begin.

Waartoe dit alles? Wordt Wit verondersteld dit allemaal in zijn geheugen op te bergen? Om de situatie scherp te doorzien is het van belang een onderscheid te maken tussen twee geheel verschillende gedragstypen, die we beide (zij het als uitersten) in de practijk tegenkomen. Eén houding komt neer op: wie dan leeft die dan zorgt, als de problemen komen zal ik ze wel oplossen maar niet voor die tijd, enz. Dit is de 'losbol' en tegenover hem staat de 'piekeraar'. Hij denkt aan mogelijke situaties in de toekomst waar geen 'normaal mens' aan zou denken. Anders gezegd, het gedrag van mensen zoals we dat normaliter observeren ligt tussen deze uitersten: men houdt wel rekening met hetgeen de toekomst zal brengen en met de gevolgen van zijn handelingen in die toekomst, maar men ziet niet onbeperkt vooruit. Het is ook duidelijk waarom. De losbol is kortzichtig en bepaald niet het ideaal; vooruitzien verdient verre de voorkeur maar men kan nu eenmaal niet alles in zich opnemen. Wat de piekeraar verweten kan worden is dat hij zoveel denkt dat hij niet tot daden komt; hij houdt er geen rekening mee dat het rendement van het denken na een zeker punt zo gering wordt, dat het verlies ten gevolge van niet-handelen belangrijker wordt. Maar stel nu eens, dat om wat voor reden dan ook het denkproces eenvoudig niet telt, dus dat we zonder merkbare moeite alle denkbare handelingen onzerzijds en alle denkbare gevolgen onder alle denkbare situaties zouden kunnen doorgronden. Dan is het kennelijk van belang dit te doen en dit is precies de situatie, die we bij ons (denkbeeldige) schaakspel op het oog hebben.

Wij keren dus weer terug naar het schaakspel. Wit doorgrondt alle denkbare mogelijkheden tot het einde toe en houdt daarbij zijn doel goed voor ogen: hij wil winnen en, mocht dit niet lukken, dan althans remise maken. Op grond van dit inzicht en deze doelstelling bepaalt hij zijn eerste zet. Daarna volgt Zwart met zijn eerste zet en onmiddellijk weet Wit wat hem te doen staat: hij heeft immers zijn tweede zet bepaald voor welke van Zwart's 20 mogelijke eerste zetten dan ook! Zo gaat het door. En dat is nu een *strategie* (ook genoemd *beslissings-*

regel): een strategie specificeert voor alle successieve stadia een beslissing, die afhankelijk is van de informatie die beschikbaar *zal zijn* op het moment dat die beslissing moet worden genomen. In ons geval: Wit's strategie stelt hem in staat iedere keer, zodra hij weer aan zet is, onmiddellijk zijn zet te doen, die dan afhankelijk is van het gehele voorgaande verloop van het spel inclusief de laatste zet van Zwart. Eenvoudig gezegd: een strategie heeft op alles zijn antwoord klaar.

Het is duidelijk, dat als Wit eenmaal een strategie heeft geformuleerd, zijn aanwezigheid bij het spel helemaal niet noodzakelijk is. Hij kan die strategie in boekvorm (een dik boek!) overhandigen aan de scheidsrechter (niet aan zijn tegenspeler – die zou er rekening mee kunnen houden!) en heengaan. Zwart zit alleen achter het bord. De scheidsrechter spoort Wit's eerste zet op uit het boek, Zwart denkt na en geeft zijn eerste zet, de scheidsrechter gaat na wat Wit hierover heeft opgeschreven, enz. Maar dit alles geldt wanneer alleen Wit een strategie heeft. Stel nu dat Zwart er ook een heeft. Dan zijn er twee boeken; beide spelers kunnen vertrekken en de scheidsrechter kan het alleen af door de boeken te vergelijken. Zelfs hij is niet meer nodig; wanneer de spelers zo behulpzaam zijn hun strategieën in kaarten te ponsen, kan het machinaal. Het spel als spel is ter ziele.

2. VOORDELIGE GELDBELEGGING

Het schaakvoorbeeld laat duidelijk zien, dat wil een strategie werkelijk 'interessant' zijn, aan een aantal voorwaarden moet zijn voldaan:

1. Er moet sprake zijn van handelingen op een aantal successieve tijdstippen.

2. De persoon in kwestie moet niet een volmaakt voorspeller zijn, d.w.z. er moeten dingen in de toekomst zijn die wel relevant zijn maar hem niet van tevoren bekend.

3. Hij moet in de loop van de tijd althans gedeeltelijke informatie over die dingen krijgen en op die informatie kunnen reageren.

Het is gemakkelijk in te zien, dat het schaakspel aan deze voorwaarden voldoet: de Witspeler moet een aantal successieve zetten doen (punt 1), hij is geen volmaakt voorspeller van de

gedragslijn van Zwart en die is voor hem bepaald relevant (punt 2) en in de loop van het spel geeft Zwart zijn zetten af en openbaart aldus die gedragslijn successievelijk, waarop Wit met zijn zetten kan reageren (punt 3). Het valt ook gemakkelijk in te zien, dat aan deze drie voorwaarden inderdaad voldaan moet zijn wil er sprake zijn van een werkelijke strategie. Wat punt 1 betreft, indien er alleen maar sprake is van een handeling op een enkel tijdstip, dan heeft het nauwelijks zin om te filosoferen over informatie die beschikbaar zal zijn op het moment dat de beslissing zal moeten worden genomen. Die ene beslissing moet dan worden genomen op basis van *de* beschikbare informatie en dat is alles. Voorts, wat punt 2 betreft, wanneer alles van tevoren bekend is komt er successievelijk geen informatie bij waarop te reageren valt. Het belang van punt 3 zal eveneens duidelijk zijn.

Onzekerheid en de geleidelijke reductie daarvan op successieve tijdstippen vormen dus belangrijke elementen van de problematiek die we hier hebben aangesneden. Nu is het van belang, dat die onzekerheid niet noodzakelijkerwijs het gevolg is van de aanwezigheid van een tegenspeler. Bij het schaken is dat wel zo; de onbekendheid met de gedragslijn van de tegenstander is daar de enige bron van onzekerheid. Maar er zijn andere situaties denkbaar, waar onzekerheid bestaat geheel buiten de spelers om. Bij de meeste kaartspelen bijv. worden de kaarten geschud en uitgedeeld; en de deelnemers verkeren in het onzekere t.a.v. wie wat heeft. In een dergelijk geval ligt het voor de hand de theorie te baseren op de waarschijnlijkheidsrekening en dat zal ook gebeuren in hetgeen volgt. Wij zullen hier nog een stap verder gaan door ons geheel te concentreren op het waarschijnlijkheidsaspect van de onzekerheid. We zullen nl. tegenspelers buiten beschouwing laten (spel van 'één persoon'); de uitbreiding komt in het volgende hoofdstuk aan de orde.

Ons voorbeeld (fictief!) is als volgt. Een Amerikaanse autofabrikant beschikt door zekere omstandigheden over een overvloed van middelen. Zijn probleem is wat daarmee te doen en wij veronderstellen dat het gaat om twee alternatieven:

1. Hij kan zijn geld in eigen branche beleggen. Helaas laat de winstgevendheid van de autobranche te wensen over.

2. Hij kan een vliegtuigfabriek overnemen. Deze richt zich i.h.b. op de productie van militaire vliegtuigen, hetgeen een voordelige zaak is.

Aannemende dat de doelstelling zuiver geldelijk van aard is, zouden wij op basis van deze gegevens concluderen dat het tweede alternatief de voorkeur verdient. Echter, de moeilijkheid is deze: op het moment is er een verkiezingscampagne gaande voor het Presidentschap van de Verenigde Staten en één van de kandidaten verkondigt, dat hij na zijn verkiezing een speciale belasting zal invoeren op winsten gemaakt op overheidscontracten. De vliegtuigfabriek valt hieronder. Wordt de belasting ingevoerd, dan is het met de winstgevendheid van die fabriek gedaan en verdient belegging in de autobranche de voorkeur. De verkiezingskansen van de kandidaat voor het Presidentschap (wij nemen aan, dat deze zich na verkiezing aan zijn woord zal houden) spelen dus een belangrijke rol in dit beslissingsprobleem.

Wij breiden het probleem met nog een stap uit door te veronderstellen, dat de autofabrikant onmiddellijk na de verkiezingen zijn positie kan herzien. Mocht hij in eerste instantie besloten hebben het op de vertrouwde autobranche te houden, dan kan hij na het bekend worden van het verkiezingsresultaat alsnog de vliegtuigfabriek kopen; en mocht hij deze fabriek in eerste instantie gekocht hebben, dan kan hij deze na de verkiezingen afstoten. En hij kan natuurlijk ook zijn oude positie continueren, d.w.z. hij kan in de autobranche blijven als hij daartoe in eerste aanleg had besloten en hij kan ook de vliegtuigfabriek behouden. Uiteraard kost dit 'switchen' geld, omdat daaraan nu eenmaal kosten verbonden plegen te zijn.

Samenvattend hebben wij dus de volgende drie kritieke momenten:

1. Nu, op dit moment, moet de autofabrikant een keuze maken uit twee alternatieven: hetzij in de autobranche blijven, hetzij de vliegtuigfabriek kopen.

2. Het moment van de verkiezingsuitslag: dan is bekend of de belasting al of niet wordt ingevoerd.

3. Onmiddellijk daarna: de autofabrikant moet opnieuw een keuze uit twee maken, nl. al of niet de vliegtuigfabriek toch kopen als hij vóór de verkiezingen besloten had in de autobranche te blijven en al of niet de vliegtuigfabriek afstoten als hij die vóór de verkiezingen had gekocht.

Merk op, dat dit de eenvoudigst denkbare situatie voor een strategiekeuze is. We hebben immers gezien, dat daartoe beslissingen op een aantal successieve tijdstippen noodzakelijk

zijn (hier hebben we het minimum, nl. beslissingen op twee tijdstippen); dat er dingen in de toekomst moeten zijn die niet met zekerheid vaststaan en van belang zijn voor het resultaat (hier: de verkiezingsuitslag); en dat hierover zekere informatie ter beschikking moet komen waarop te reageren valt (hier: door 'switchen'). Aan de drie voorwaarden is dus voldaan en wel inderdaad op een manier die zo eenvoudig mogelijk is: bij beide beslissingen gaat het om een keus uit slechts twee alternatieven en bovendien is de grote onbekende van het probleem – de verkiezingsuitslag – van hetzelfde type, nl. één mogelijkheid uit twee.

3. DE WAARDERING VAN DE RESULTATEN

Wij hebben zojuist gezien, dat elk der drie momenten gekenmerkt is door twee alternatieven. In totaal zijn er dus $2 \times 2 \times 2 = 8$ verschillende mogelijkheden. Bijvoorbeeld:

De fabrikant koopt de vliegtuigfabriek vóór de verkiezingen;
de kandidaat in kwestie wordt gekozen;
de fabrikant doet de vliegtuigfabriek na de verkiezingen van de hand.

Of ook:

De fabrikant besluit vóór de verkiezingen de vliegtuigfabriek niet te kopen;
de kandidaat wordt gekozen;
de fabrikant besluit ook na de verkiezingen in de autobranche te blijven.

En zo nog zes andere.

De fabrikant dient nu zijn waardering voor alle acht mogelijkheden numeriek vast te stellen. In ons voorbeeld gaan wij daartoe als volgt te werk. Neem eerst het geval, waarin de fabrikant vòòr de verkiezingen besluit om in de autobranche te blijven en in deze houding na de verkiezing volhardt. Dit geval omvat nog steeds twee mogelijkheden, want de kandidaat kan al of niet gekozen worden. Wij zullen ervan uitgaan, dat de op-

brengst in beide gevallen dezelfde is; dit ligt in de rede, want de fabrikant houdt zich dan buiten de vliegtuigbranche, terwijl de autobranche niet door de belasting wordt getroffen. (Hij maakt personenauto's, niet voor het leger.) Laat die opbrengst 2 zijn, gemeten in zekere eenheden, bijv. miljoenen dollars. Dit zal aldus worden aangegeven:

AbA: 2 *AnA:* 2.

Hierbij geeft de eerste *A* van *AbA* aan, dat de fabrikant vòòr de verkiezingen besluit in de autobranche (*A* = auto) te blijven. De *A* op het eind geeft aan, dat de fabrikant na de verkiezingen besluit in de autobranche te blijven. De *b* ertussen geeft aan, dat de belasting (*b* = belasting) wordt geheven; daarentegen geeft de *n* van *AnA* aan, dat die belasting niet (*n* = niet) wordt geheven. Aldus wordt op beknopte wijze met drie symbolen het verloop van het proces op de drie successieve momenten beschreven.

Wij bezien nu het geval, waarin de fabrikant vòòr de verkiezingen besluit de vliegtuigfabriek te kopen en na de verkiezingen die fabriek te behouden. Ook hier geldt, dat dit geval twee mogelijkheden omvat afhankelijk van de vraag of de kandidaat het bij de verkiezingen haalt of niet; maar nu is het natuurlijk niet langer waar, dat beide mogelijkheden tot dezelfde opbrengst leiden. Wij zullen ervan uitgaan, dat de opbrengst 5 is als de belasting er niet komt maar —5 (dus een verlies) als die er wel komt. Dit wordt als volgt genoteerd:

VbV: —5 *VnV:* 5.

Hierbij geeft, analoog aan het voorgaande, de eerste *V* van *VbV* de beslissing aan van vòòr de verkiezingen om de vliegtuigfabriek (*V* = vliegtuig) te kopen, de laatste *V* de beslissing van na de verkiezingen om die fabriek te houden, enz. We zien dus, dat als de belasting niet wordt opgelegd het voordeliger is het op de vliegtuigfabriek te houden (5 voor *VnV* tegen 2 voor *AnA*); maar ook, dat als de belasting wèl wordt opgelegd, het vèèl *on*voordeliger wordt de vliegtuigfabriek te kopen (—5 voor *VbV* tegen 2 voor *AbA*).

We zijn er echter nog niet, omdat we ons bepaald hebben tot de gevallen, waarin de fabrikant zijn positie na de verkiezingen niet wijzigt. Bezien we dus nu de gevallen van 'switchen'. Hun opbrengsten zullen worden afgeleid uit de voorgaande aan de hand van de twee volgende veronderstellingen: (i) wat er ook gebeurt, de opbrengst wordt volledig bepaald door de uitslag

van de verkiezingen en door de positie die de fabrikant onmiddellijk daarna kiest, echter (ii) van deze opbrengst moet 1 worden afgetrokken voor de aan het 'switchen' verbonden kosten. Als voorbeeld nemen we *VbA*, hetgeen betekent: de fabrikant besluit vòòr de verkiezingen de vliegtuigfabriek te kopen (*V*), de kandidaat wordt gekozen (*b*) en de fabrikant doet daarna de vliegtuigfabriek van de hand (*A*). Uit veronderstelling (i) volgt, dat de opbrengst volledig bepaald wordt door 'de kandidaat wordt gekozen en de fabrikant doet daarna de vliegtuigfabriek van de hand'. Maar *AbA* heeft precies dezelfde eigenschap: ook voor die variant geldt, dat de kandidaat wordt gekozen en dat de fabrikant het verder zonder vliegtuigfabriek doet. Beide varianten hebben dus dezelfde opbrengst, echter voor *VbA* moet 1 worden afgetrokken omdat de fabrikant zijn positie heeft herzien (van *V* naar *A*). Aangezien de opbrengst van *AbA* hierboven vastgesteld is op 2, wordt die van *VbA* gelijk 1. Voor *VnA* gaat het op dezelfde manier. De gedragslijn van de fabrikant is precies dezelfde als in het geval *VbA*, maar nu komt het zo uit, dat de kandidaat niet is gekozen. De opbrengst van *VnA* wordt verkregen door die van *AnA* met 1 te verminderen, dus ook 1. Samenvattend:

VbA: 1 *VnA:* 1.

Tenslotte hebben we *AbV* en *AnV*: de fabrikant houdt het vòòr de verkiezingen op de autobranche, de kandidaat wordt wel resp. niet gekozen, en na de verkiezingen koopt hij de vliegtuigfabriek. Evenals in de twee voorgaande gevallen wordt de opbrengst bepaald door de laatste twee letters (*bV* en *nV*), want die geven de uitslag van de verkiezingen aan alsmede de gekozen positie daarna. [Zie veronderstelling (i).] Wij vinden dus de opbrengsten van *AbV* en *AnV* door 1 af te trekken (de fabrikant 'switcht' van *A* naar *V*!) van *VbV* resp. *VnV* en het resultaat is:

AbV: —6 *AnV:* 4.

De opbrengst —6 van *AbV* is de laagste die er bij is, maar dat ligt dan ook voor de hand! Eerst de vliegtuigfabriek afwijzen zolang er nog hoop was dat die fabriek een winstgevende zaak zal zijn, en hem vervolgens toch kopen wanneer aan die hoop iedere grond is ontnomen is een gedragslijn, die niet aantrekkelijk kan zijn. Toch zullen we *AbV* in onze beschouwingen betrekken; we zijn immers volgens de eerste bladzijden van dit hoofdstuk verplicht systematisch te werk te gaan.

Wij hebben nu alle 8 mogelijkheden bezien, waarvan we de opbrengsten nog even samenvatten:

AbA:	2	*AnA:*	2
VbV:	—5	*VnV:*	5
VbA:	1	*VnA:*	1
AbV:	—6	*AnV:*	4.

Dit is een *betaalrooster* (de Nederlandse vertaling van 'pay-off matrix', afkomstig van de Utrechtse hoogleraar Freudenthal), d.i. een overzicht van de opbrengst bij alle denkbare gedragslijnen en onder alle denkbare omstandigheden. De vraag is dan nu: wat dient de autofabrikant in concreto te doen? Hij wil een zo hoog mogelijke opbrengst en zal constateren dat de hoogste die van *VnV* is, nl. 5. Dit suggereert dat hij de vliegtuigfabriek nu (dus vòòr de verkiezingen) dient te kopen; immers, dàt is de betekenis van de eerste *V* van *VnV*. Echter, de opbrengst van *VnV* is daarom zo hoog, omdat ervan is uitgegaan dat de kandidaat in kwestie niet wordt gekozen. Maar dat heeft onze fabrikant niet in de hand. Loopt het mis, d.w.z. wordt de kandidaat wel gekozen, dan belandt hij bij *VbV* en zijn opbrengst wordt —5, dus bepaald onaantrekkelijk. Hij zou aldus kunnen redeneren: ik kies die gedragslijn die mij in het ongunstigste geval nog de grootste winst oplevert. Welnu, dan dient hij niet naar de vliegtuigfabriek om te kijken. De opbrengsten van*AbA* en *AnA* zijn immers beide 2, dus door in de autobranche te blijven, zowel vòòr als na de verkiezingen, is hij zeker van een winst van 2; en zoals we zojuist zagen, loopt hij het risico van —5 in geval hij de vliegtuigfabriek koopt en aanhoudt (*VbV* en *VnV*), terwijl bij 'switchen' de laagste winst eveneens beneden 2 is: 1 voor *VbA* en *VnA*, —6 voor *AbV* en *AnV*.

Het criterium van de hoogste opbrengst in het ongunstigste geval hebben we gehanteerd bij het wedden op Duindigt (blz. 37). Daarna zijn we op dit criterium teruggekomen toen de waarschijnlijkheidsrekening werd geïntroduceerd (blz. 158); we vroegen ons toen af of het criterium wel gebruik maakte van onze kennis van zaken betreffende paardenraces. En evenzo hier: stel bijv. dat de fabrikant er bijna zeker van is, dat de kandidaat niet wordt verkozen. In dat geval wordt de vliegtuigfabriek toch wel veel aantrekkelijker, want de gun-

stige opbrengst van *VnV* heeft dan aanmerkelijk meer kans om gerealiseerd te worden dan de ongunstige opbrengst van *VbV*.

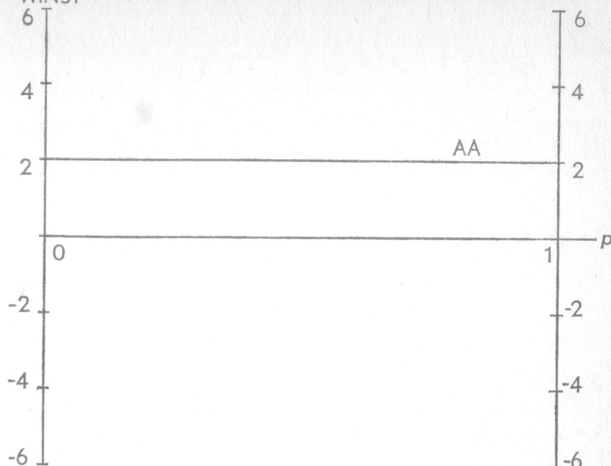
Het criterium, dat hier gehanteerd zal worden, komt op het volgende neer. Laat de autofabrikant de waarschijnlijkheid, dat de kandidaat wordt gekozen, op p stellen, enig getal tussen 0 en 1. Dan kiest hij de strategie waarvoor de mathematische verwachting van de opbrengst het hoogst is, gegeven de numerieke opbrengstwaarden van de hierboven beschouwde 8 mogelijkheden en gegeven een zekere numerieke waarde van p.

4. VIER PAREN VAN OPEENVOLGENDE BESLISSINGEN

Wij gaan in eerste aanleg als volgt te werk. De fabrikant moet nù beslissen of hij de vliegtuigfabriek al of niet koopt en nà de verkiezingen of hij die fabriek behoudt of van de hand doet resp. alsnog koopt of niet koopt. Dan zijn de volgende vier gedragslijnen denkbaar:

AA: de fabrikant blijft nu in de autobranche en continueert deze houding na de verkiezingen.
AV: de fabrikant blijft nu in de autobranche en koopt de vliegtuigfabriek na de verkiezingen.
VA: de fabrikant koopt nu de vliegtuigfabriek en doet die na de verkiezingen van de hand.
VV: de fabrikant koopt nu de vliegtuigfabriek en behoudt deze na de verkiezingen.

Voor elk van dit viertal zullen wij nu de verwachte winst bepalen. Bij *AA* is dit heel eenvoudig. Deze gedragslijn komt neer op de mogelijkheid *AbA* wanneer de kandidaat wordt verkozen, op *AnA* wanneer deze niet wordt verkozen. In beide gevallen is de winst 2 en dus ook is de winst van *AA* gelijk 2 (en wel met zekerheid). Anders gezegd, hoe groot de kans p dat de kandidaat wordt gekozen ook is, de winst van *AA* is altijd 2; en dus ook is de mathematische verwachting van de winst gelijk aan 2. Het zal nuttig blijken dit grafisch af te beelden. Daartoe zetten we langs de horizontale as de kans p af; verticaal meten we de verwachte winst. Deze wordt voor *AA* weergegeven door een horizontale rechte op de hoogte 2.

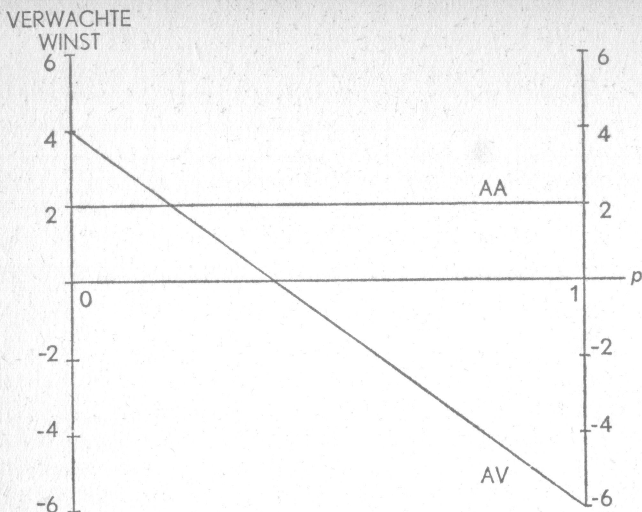


In deze grafiek kan de horizontaal afgezette variabele (p) geen grotere waarde dan 1 aannemen, vandaar de tweede verticale as bij $p = 1$, die de figuur afsluit. Deze extra as zal ons de mogelijkheid geven onze verdere rechte lijnen wat gemakkelijker af te zetten.

Vervolgens bezien we *AV*. Deze gedragslijn komt met kans p neer op de mogelijkheid *AbV* (de kandidaat wordt gekozen) en met kans $1 - p$ op *AnV* (de kandidaat wordt niet gekozen). De eerste mogelijkheid heeft een opbrengst —6, de tweede 4, zie het overzicht op blz. 193; dus is de verwachting van de opbrengst van *AV*:

$$p(-6) + (1 - p)4 = 4 - 10p.$$

Deze verwachte opbrengst is kennelijk *niet* onafhankelijk van de kans dat de kandidaat wordt gekozen. Is die kans 0, dan is de verwachte opbrengst 4; is de kans 1, dan is de verwachte opbrengst —6; en voor tussenliggende waarden daalt de verwachte opbrengst lineair bij stijgende p. Wij krijgen dus nu een rechte lijn, die begint bij 4, dan daalt en eindigt bij —6. Op blz. 196 is deze lijn aan de vorige figuur toegevoegd. Uit de resulterende figuur blijkt direct dat *AV* een grotere verwachte opbrengst heeft dan *AA* zolang de kans p voldoende klein is, maar een lagere wanneer p groot is. Dat ligt in de rede: *AV* betekent

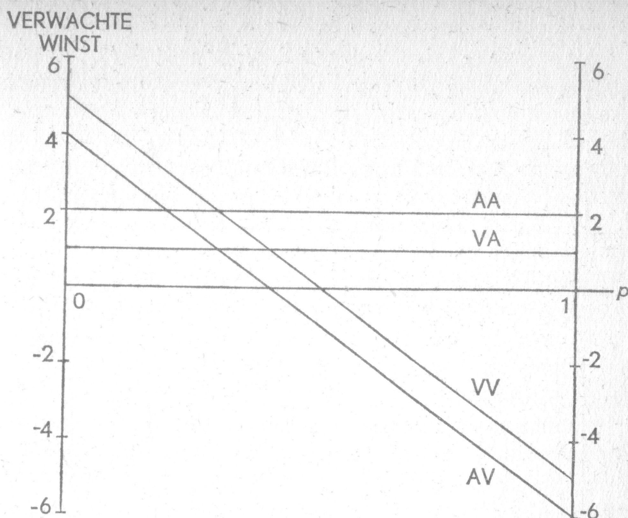


immers, dat de fabrikant na de verkiezingen de vliegtuigfabriek bezit, voor *AA* geldt het tegendeel, en het bezit van die vliegtuigfabriek is slechts voordelig wanneer de kandidaat de verkiezingen niet haalt. Welnu, een kleine waarde van p betekent voor de fabrikant, dat de kandidaat de verkiezingen vermoedelijk niet zal halen; de verwachte opbrengst van *AV* stijgt dienovereenkomstig boven het niveau van *AA*. Het tegendeel geldt voor een hoge waarde van p.

Tenslotte *VA* en *VV*. Wat *VA* betreft, de verwachte opbrengst is hier heel eenvoudig te bepalen. Met kans p komt deze gedragslijn neer op *VbA* (met opbrengst 1) en met kans $1 - p$ op *VnA* (ook met opbrengst 1); de opbrengst van *VA* is dus 1 met zekerheid en wordt weergegeven door een horizontale rechte op de hoogte 1. Voor *VV* geldt, dat de opbrengst —5 is met kans p (de mogelijkheid *VbV*) en 5 met kans $1 - p$ (*VnV*), dus is zijn verwachte opbrengst:

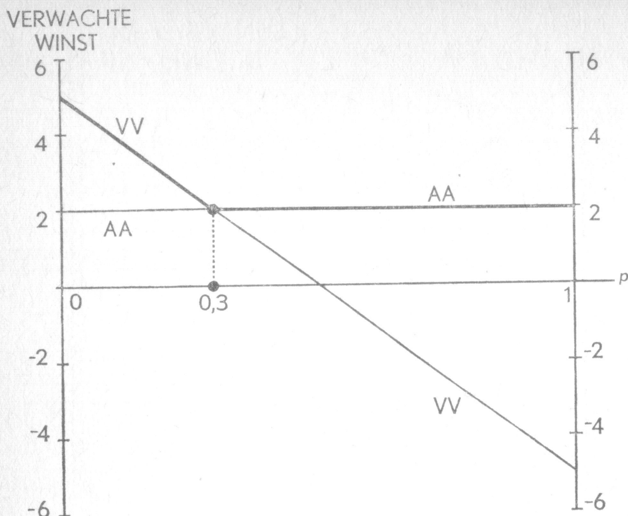
$$p(-5) + (1 - p)5 = 5 - 10p.$$

De lijn is nu dalend, begint op niveau 5 bij $p = 0$ en eindigt op —5 bij $p = 1$. De volledige figuur treft u aan op blz. 197. Wat is nu de beste gedragslijn? Het antwoord ligt voor de hand: dit hangt ervan af, hoe groot p is. Dit is juist, echter met een kleine kanttekening. We zien nl., dat de verwachte opbrengst



van *VA* altijd beneden *AA* ligt en evenzo voor *AV* t.o.v. *VV*, hoe groot *p* ook is. Er is dus geen enkele omstandigheid denkbaar, waaronder *VA* de voorkeur verdient boven *AA* of *AV* boven *VV*. Dit drukt men uit door te zeggen, dat *VA* en *AV* 'gedomineerd' worden door *AA* resp. *VV*. Wij zullen daarom *VA* en *AV* (de twee gevallen van 'switch') buiten beschouwing laten; die komen kennelijk toch niet in aanmerking. Laat men de bijbehorende lijnen weg dan wordt de figuur aanzienlijk vereenvoudigd, zoals men ziet op de volgende bladzijde bovenaan.

In deze figuur zijn twee lijnstukken dik getekend. Deze geven voor alle *p*-waarden de *hoogste* verwachte opbrengst aan. Voor kleine *p* is dit de verwachte opbrengst van *VV*, voor grote *p* die van *AA*. Waar ligt precies het punt van overgang? Om deze vraag te beantwoorden roepen wij in herinnering, dat de verwachte opbrengst van *VV* gelijk is aan $5 - 10p$ en die van *AA* aan 2. Nu is $5 - 10p$ groter dan 2 zolang *p* kleiner is dan 0,3 en kleiner dan 2 zodra *p* boven 0,3 uitkomt. Dus verdient *VV* de voorkeur boven *AA* als *p* kleiner is dan 0,3 (en andersom als *p* groter is). Maar *VV* impliceert, dat de autofabrikant op dit moment de vliegtuigfabriek koopt en *AA* dat hij dit niet doet. De conclusie luidt: de autofabrikant dient nu, op dit moment, de vliegtuigfabriek te kopen als hij de kans dat de



kandidaat wordt gekozen lager stelt dan 0,3; hij dient niet te kopen als die kans boven 0,3 gesteld wordt.

Deze conclusie is foutief! Wij zullen nu zien waarom.

5. VIER STRATEGIEËN

De conclusie is foutief omdat wij de fabrikant geen rekening hebben laten houden met (i) de omstandigheid dat hij op het moment van de verkiezingsuitslag weet dat de belasting er komt of niet komt en (ii) met het feit dat hij in de gelegenheid is op dat moment van die kennis gebruik te maken door zijn positie te herzien. Anders gezegd, wij hebben geen werkelijke strategieën bekeken, die immers van tevoren specificeren hoe gereageerd zal worden op alle denkbare eventualiteiten.

Hoe ziet zo'n strategie er dan uit? Welnu, de volgende is een voorbeeld:

De fabrikant koopt de vliegtuigfabriek vóór de verkiezingen.
Als de kandidaat wordt gekozen doet hij de fabriek weer van de hand.
Als de kandidaat niet wordt gekozen behoudt hij de fabriek.

We zien hier duidelijk, dat de beslissing voor de tweede periode niet onvoorwaardelijk wordt gespecificeerd, zoals in de hiervoor beschreven gevallen *AA*, *AV*, *VA* en *VV;* integendeel, die tweede beslissing wordt afhankelijk gemaakt van de informatie waarover de fabrikant zal beschikken op het moment, waarop hij die tweede beslissing moet nemen. Dit is zoals het een rechtgeaarde strategie betaamt. Wij zullen hem aangeven met *V*(*bA; nV*). Hierbij representeert de *V* vòòr de haken de beslissing van vòòr de verkiezingen om de vliegtuigfabriek te kopen. De symbolen tussen haken slaan op de beslissing van na de verkiezingen. Vòòr de puntkomma staat *bA*, hetgeen wil zeggen: als de kandidaat het haalt en dus de belasting (*b*) er komt, dan terug naar de autobranche (*A*). Na de puntkomma staat *nV* met de volgende betekenis: als de kandidaat het niet (*n*) haalt, dan wordt de vliegtuigfabriek (*V*) behouden.

Dit is natuurlijk niet de enige strategie. Een andere is *V*(*bV; nA*), waarbij de beslissing van vòòr de verkiezingen dezelfde is, maar daarna precies andersom gehandeld wordt: komt de belasting (*b*), dan de vliegtuigfabriek (*V*) behouden; komt hij er niet (*n*), dan terug naar de autobranche (*A*). Dat is natuurlijk geen bijzonder intelligente strategie.

Er zijn nog twee andere strategieën, nl. *A*(*bA; nV*) en *A*(*bV; nA*). In beide gevallen besluit de fabrikant vòòr de verkiezingen in de autobranche te blijven; in het eerste geval wordt deze positie gecontinueerd als de kandidaat wordt gekozen maar de vliegtuigfabriek gekocht als hij niet wordt gekozen, in het tweede geval is het juist andersom. Dat er in ons probleem niet meer dan vier strategieën zijn valt gemakkelijk in te zien. Er zijn nl. 2 mogelijkheden voor een beslissing vóór de verkiezingen, *A* en *V*; er zijn ook 2 mogelijkheden om na de verkiezingen op de uitslag te reageren, nl. (*bA; nV*) en (*bV; nA*). In totaal dus $2 \times 2 = 4$. Merk op, dat een strategie van het type *A*(*bV; nV*) niet als een serieuze strategie wordt meegeteld. Dan wordt immers, of de belasting er komt (*b*) of niet (*n*), steeds de vliegtuigfabriek (*V*) gekocht en dus is de beslissing van na de verkiezingen onafhankelijk van de uitslag. Gemakkelijk valt in te zien, dat *A*(*bV; nV*) identiek is met *AV*; en deze gedragslijn hebben we hierboven al onder ogen genomen.

De lezer zal zich wellicht afvragen: Dit alles is een spitsvondige verfijning, wellicht nuttig om het probleem nog iets dieper te doorgronden, maar maakt het iets uit voor de beslissing die

vòòr de verkiezingen moet worden genomen? Wat de fabrikant in feite te doen heeft komt neer op: nu al of niet de vliegtuigfabriek kopen en na de verkiezingen die fabriek al of niet behouden respectievelijk al of niet toch kopen. Dus komt zijn uiteindelijke gedrag neer op één van de vier alternatieven *VV*, *VA*, *AV*, *AA*. En de verwachte opbrengst – het criterium – van dit viertal hebben we vergeleken. Wat valt er dan nog verder te bezien?

Het volgende. Om de verwachte opbrengst van *A*(*bA*; *nV*) te berekenen dienen wij ons te realiseren, dat deze strategie met kans p neerkomt op de mogelijkheid *AbA* en met kans $1 - p$ op *AnV*. Immers, de strategie houdt in: eerst *A* vòòr de verkiezingen en dan wanneer *b* zich realiseert (de kandidaat haalt het – en de kans daarop is p), ook *A* na de verkiezingen; maar als *n* zich realiseert (de kandidaat haalt het niet – de kans daarop is $1 - p$), dan *V* na de verkiezingen. Dus is de verwachte opbrengst van deze strategie gelijk aan p maal de opbrengst van *AbA* plus $1 - p$ maal de opbrengst van *AnV*, dus

$$p \times 2 + (1 - p)4 = 4 - 2p,$$

zie het overzicht op blz. 193. Voor de drie overige strategieën gaan we op analoge wijze te werk. Men ziet gemakkelijk in, dat de verwachte opbrengsten van

A(*bV*; *nA*), *V*(*bA*; *nV*), *V*(*bV*; *nA*)

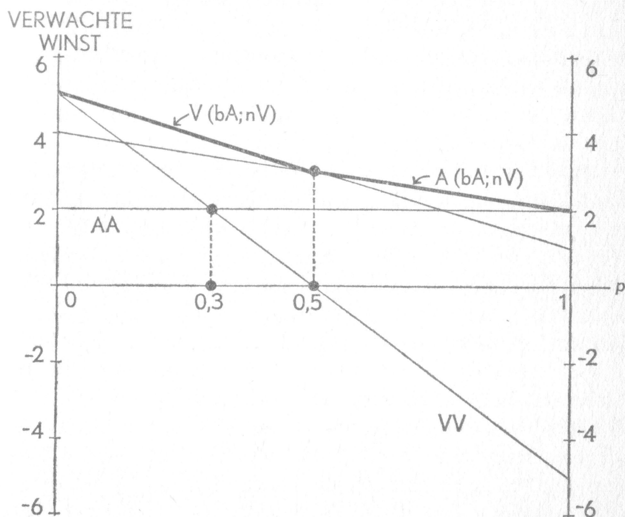
gevonden worden door p maal de opbrengst van resp. *AbV*, *VbA*, *VbV* te nemen en daarbij $1 - p$ maal de opbrengst van resp. *AnA*, *VnV*, *VnA* op te tellen. Aldus vinden we de volgende verwachte opbrengsten van de vier strategieën:

A(*bA*; *nV*): $4 - 2p$
A(*bV*; *nA*): $2 - 8p$
V(*bA*; *nV*): $5 - 4p$
V(*bV*; *nA*): $1 - 6p$

We brengen in herinnering (zie blz. 194), dat de gedragslijn *AA* een opbrengst van 2 garandeert. Dit betekent, dat *A*(*bV*; *nA*) en *V*(*bV*; *nA*) door deze gedragslijn gedomineerd worden; hun verwachte opbrengsten liggen immers altijd beneden 2, hoe groot of hoe klein de kans van de kandidaat ook is.[1] Echter, *AA* wordt op zijn beurt gedomineerd door *A*(*bA*; *nV*), waarvan

1. Behalve wanneer $p = 0$, want dan hebben *AA* en *A* (*bV*; *nA*) dezelfde verwachte opbrengst (2). Maar als $p = 0$, dan is het hele probleem de wereld uit, want dan is er geen onzekerheid! Hetzelfde geldt voor $p = 1$.

de verwachte opbrengst steeds boven 2 ligt. Bovendien zien we, dat *V(bA; nV)* een verwachte opbrengst heeft hoger dan die van *VV*, die immers 5 — 10*p* is (zie blz. 196). Nu waren *AA* en *VV* de enige gedragslijnen die in aanmerking kwamen, zolang we nog geen werkelijke strategieën bekeken; kennelijk worden zij elk door een werkelijke strategie gedomineerd. Het is instructief dit in een figuur af te beelden:



De dik getekende lijnstukken liggen nu hoger dan eerst het geval was; bovendien ligt het punt van overgang meer naar rechts. Voor kleine *p* is de strategie *V(bA; nV)* het voordeligst, die in woorden op het volgende neerkomt:

> De fabrikant koopt de vliegtuigfabriek vòòr de verkiezingen. Als de kandidaat wordt gekozen doet hij de fabriek van de hand.
> Als de kandidaat niet wordt gekozen behoudt hij de fabriek.

Voor grote *p* wint *A(bA; nV)* het, dus:

> De fabrikant besluit vòòr de verkiezingen in de autobranche te blijven.

Als de kandidaat wordt gekozen blijft hij ook daarna in de autobranche.
Als de kandidaat niet wordt gekozen koopt hij de fabriek alsnog.

Dit tweetal strategieën vervangt *AA* en *VV*. Maakt het ook iets uit voor de beslissing, die vòòr de verkiezingen moet worden genomen? Zeker! De eerste strategie heeft als verwachte opbrengst 5 — $4p$, de tweede 4 — $2p$. Zolang p kleiner is dan 0,5 is de eerste waarde groter dan de tweede; is p groter dan 0,5, dan geldt het omgekeerde. De eerste strategie impliceert de beslissing van het kopen van de vliegtuigfabriek vòòr de verkiezingen, de tweede het tegendeel. Dus is de kritieke waarde van p gestegen van 0,3 tot 0,5! Anders gezegd: eerst concludeerden we, dat boven een kans van 0,3 van de kandidaat de fabrikant in elk geval tot de verkiezingen in de autobranche moet blijven; nu zien we dat dit onjuist is en dat hij deze beslissing pas dient te nemen als de kans van de kandidaat boven 0,5 ligt.

Wat is de reden van deze gewijzigde conclusie? Het feit dat we strategieën bezien hebben. Kan dit resultaat ook intuïtief plausibel worden gemaakt? Dat kan. Ons resultaat komt hierop neer, dat de inschakeling van strategieën de aantrekkelijkheid van de vliegtuigfabriek verhoogt. Zolang we ons bepaalden tot gedragslijnen, waarbij de beslissing van na de verkiezingen niet afhankelijk van de verkiezingsuitslag wordt gemaakt, was het *VV* die als enige op grond van verwachte winstgevendheid de aankoop van de vliegtuigfabriek impliceert. Deze gedragslijn (*VV*) onderschat de winstgevendheid van die beslissing. Hij gaat er nl. van uit, dat die fabriek behouden wordt ook als de kandidaat wordt gekozen, in welk geval een verlies het resultaat is. Hij ziet over het hoofd, dat er een manier is om dit verlies te vermijden, nl. afstoten. En dit wordt *niet* over het hoofd gezien door de strategie *V*(*bA*; *nV*), die gekenmerkt is door dezèlfde beslissing van vòòr de verkiezingen, nl. *V*. En door al van tevoren rekening te houden met de mogelijkheid van afstoten (hetgeen de strategie doet, maar niet *VV*) komt de fabrikant tot een optimistischer (en juister!) beeld van de merites van de vliegtuigfabriek. En dit verhoogt de kritieke waarde van de kans van de kandidaat, beneden welke het voordelig wordt die fabriek nu, op dit moment, te kopen. Aldus zien we, dat het rekening houden

met het feit dat men in de toekomst informatie zal ontvangen
en dat men op die informatie zal kunnen reageren,
van invloed kan zijn op het nemen van rationele beslissingen in het heden.

LITERATUUR

Het strategiebegrip is een essentieel onderdeel van de speltheorie, die in het volgende hoofdstuk ter sprake komt. Het standaardwerk is dat van Von Neumann en Morgenstern [1], dat echter verre van eenvoudig is. Het hier gekozen voorbeeld is afkomstig van Hoofdstuk 8 van Theil [2]. Voor andere toepassingen verwijzen we naar Schlaifer [3] en Luce and Raiffa [4].

[1] Neumann, J. von, and O. Morgenstern, *Theory of Games and Economic Behavior*. Third Edition. Princeton University Press, Princeton. 1953.

[2] Theil, H., *Economic Forecasts and Policy*. Second Edition. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam. 1961.

[3] Schlaifer, R., *Probability and Statistics for Business Decisions*. McGraw-Hill Book Company, Inc., New York. 1959.

[4] Luce, R. D., and H. Raiffa, *Games and Decisions*. John Wiley and Sons, Inc., New York. 1957.

Nog een kanttekening over het maximeren van de verwachte opbrengst – de in dit hoofdstuk gevolgde procedure. Dat dit een beperking inhoudt wordt duidelijk uit het volgende voorbeeld. Stel dat U in de gelegenheid bent om een lot te kopen voor de volgende loterij. Er wordt een munt opgegooid, en als er kruis boven komt ontvangt U niets, maar komt er munt boven dan ontvangt U 100.000 gulden. De verwachte opbrengst is dus $^1/_2 \times 0 + ^1/_2 \times 100.000 =$ 50.000 gulden. Stel dat U het lot kunt kopen voor 49.000 gulden. De verwachte winst is dan 1000 gulden. Het alternatief is dat U het lot niet koopt. U ontvangt dan niets, de verwachte winst is nul. Wanneer U de verwachte winst zou maximeren, zou U het lot dus kopen. Vraag U nu af of U dit werkelijk zou doen, als U voor de keus gesteld werd. U riskeert 49.000 gulden te verliezen. De meeste mensen willen dit risico stellig niet lopen, en kopen dus het lot niet. Zij maximeren dus de verwachte winst niet. Toch wordt de theorie door dit bezwaar niet zozeer aangetast als U wellicht denkt. Er kan nl. (onder zekere, overigens tamelijk plausibele veronderstellingen), aangetoond worden, dat er steeds een zekere functie van de opbrengst bestaat, waarvan de verwachting wèl gemaximeerd wordt. Die functie kan bijv. de logarithme van de opbrengst zijn. De theorie is dan weer op de functie (niet op de opbrengst zelf) toepasbaar.

8. SPELTHEORIE

1. TWEE LUCHTVAARTMAATSCHAPPIJEN

Twee luchtvaartmaatschappijen, *A* en *B*, bedienen hetzelfde traject. Beide maatschappijen zijn erop uit zoveel mogelijk van de markt te veroveren. Het totale vervoer van passagiers per jaar op dat traject is gegeven; dus wat de één wint aan passagiers verliest de ander. We nemen aan dat drie gedragslijnen ter beschikking staan:

(1) niets doen;
(2) films vertonen tijdens de vlucht;
(3) adverteren in dag- en weekbladen.

Deze drie mogelijkheden staan zowel voor *A* als voor *B* open. Ze kunnen dus beide niets doen of beide films vertonen; het is ook mogelijk dat *A* niets doet en *B* gaat adverteren, enz. We gaan er gemakshalve van uit, dat geen twee dingen (films en advertenties) tegelijk gedaan kunnen worden. In totaal zijn er $3 \times 3 = 9$ mogelijkheden. Elk daarvan heeft zekere gevolgen voor de verdeling van de markt. Doen beide niets, dan blijft alles bij het oude. Gaat *A* adverteren terwijl *B* niets doet, dan stijgt het aantal passagiers per jaar bij *A* met 350 en het aantal bij *B* daalt met 350. (Het *totale* aantal passagiers verandert niet!) Gaat *B* adverteren terwijl *A* niets doet, dan stijgt het aantal passagiers bij *B* met 300 en bij *A* daalt het met hetzelfde aantal. Dit effect is geringer dan in het tegenovergestelde geval (300 tegenover 350); blijkbaar hebben de reclamespecialisten bij *A* er meer slag van dan die bij *B*.

Zo gaan we door voor alle 9 mogelijkheden. Aangezien de passagierswinsten van *A* en *B* steeds tegengesteld gelijk zijn, kunnen we ons gevoegelijk bepalen tot de winst van *A* (die uiteraard negatief kan zijn). De volgende tabel van *A*'s passagierswinst geeft een volledig overzicht:

Gedragslijn van *A*	Gedragslijn van *B* 1 (niets)	2 (films)	3 (adverteren)
1 (niets)	0	—100	—300
2 (films)	100	0	—100
3 (adverteren)	350	125	50

De vraag is nu: gegeven dat er voor elk drie gedragslijnen zijn en gegeven het zojuist gepresenteerde betaalrooster, wat valt er dan te zeggen over de gedragslijn die voor *A* het beste is? En voor *B*? In dit geval is het antwoord eenvoudig genoeg. Gaat het nl. uitsluitend om het winnen van passagiers en laten we dus de kosten verbonden aan de diverse gedragslijnen buiten beschouwing, dan kan *A* het beste gaan adverteren. Wat *B* ook doet, die gedragslijn geeft *A* altijd de hoogste winst: 350 tegen 0 en 100 als *B* niets doet, 125 tegen —100 en 0 als *B* films gaat vertonen, 50 tegen —300 en —100 als *B* ook gaat adverteren. En die laatste gedragslijn is ook voor *B* het beste. Hij heeft immers steeds de grootste winst of het kleinste verlies door zelf ook te adverteren: een winst van 300 als *A* niets doet en van 100 als *A* films gaat vertonen en een verlies van 50 als *A* gaat adverteren. (Dit zijn de getallen van de laatste kolom, echter met omgekeerd teken omdat de winst van *B* minus de winst van *A* is.) De conclusie is dus: *A* en *B* dienen beide te adverteren en deze combinatie van gedragslijnen leidt tot een winst van 50 passagiers per jaar van *A*, dus tot eenzelfde verlies voor *B*.

In de meeste gevallen is er geen gedragslijn die, zoals hier het adverteren, alle andere gedragslijnen domineert; d.w.z. tot een hogere opbrengst leidt onafhankelijk van wat de tegenstander doet. Het probleem wordt dan veel moeilijker. De *speltheorie* houdt zich met dit soort situaties bezig; deze theorie is ruim 20 jaren geleden ontwikkeld door de mathematicus John von Neumann en de econoom Oskar Morgenstern. Een aantal belangrijke elementen van deze theorie kwam in ons voorbeeld al naar voren: er is een aantal 'spelers' (niet noodzakelijk twee), ze hebben een aantal alternatieve gedragslijnen ter beschikking (niet noodzakelijk hetzelfde aantal voor alle spelers) en er moet gespecificeerd worden wat het resultaat is voor elk der spelers wanneer ieder zijn keuze t.a.v. de concrete gedragslijn heeft gemaakt (het betaalrooster).[1] In hetgeen volgt zullen we ons in hoofdzaak bepalen tot spelen met twee personen en voorts tot het geval van zgn. nulspelen, d.w.z. spelen waarbij de een wint wat de ander verliest, dus waarbij de som

1. Eventueel is er ook nog een waarschijnlijkheidsmechanisme, dat zijn invloed op de uitkomst van het spel doet gelden, geheel los van de gedragingen van de spelers. Het beleggingsvoorbeeld van het vorige hoofdstuk (met slechts één speler) valt onder deze categorie.

van de opbrengsten van de twee spelers nul is. Ons voorbeeld voldoet aan beide voorwaarden. Deze beperking maken we niet omdat dit in de practijk het belangrijkste geval zou zijn. Integendeel, de belangenstrijd die we in feite waarnemen kent gewoonlijk meer spelers dan twee; bovendien is het meestal niet waar, dat winsten en verliezen elkaar opheffen. De beperking wordt enerzijds gemaakt omdat de algemenere theorie bijzonder ingewikkeld is, anderzijds omdat die theorie bovendien nog in een verre van volmaakt stadium verkeert. Aan het einde van dit hoofdstuk zullen wij summier op deze problemen ingaan.

2. SPEL, ZET EN STRATEGIE

Eerst een heel eenvoudig spel om te laten zien, dat er lang niet altijd dominerende gedragslijnen bestaan: *kruis of munt*. Er zijn twee spelers, *A* en *B*, elk met een zilveren gulden. Zij dienen elk, naar eigen keuze (maar zonder te zien wat de ander doet), hun gulden met kruis dan wel munt naar boven op tafel te leggen. Dat is alles. Indien dan blijkt dat beide guldens kruis tonen of beide munt, dan krijgt *A* de gulden van *B*. Zijn ze verschillend, dan krijgt *B* de gulden van *A*.

De spelers hebben dus elk twee alternatieve gedragslijnen: kruis of munt. Het is een nulspel, want wat de een wint verliest de ander. Wat *A* ontvangt (in guldens) in de $2 \times 2 = 4$ verschillende situaties blijkt uit het volgende betaalrooster:

Gedragslijn van *A*	Gedragslijn van *B* kruis	munt
kruis	1	—1
munt	—1	1

Kennelijk is het voor *A* beter kruis boven te leggen als *B* ook kruis boven legt (zie de eerste kolom) maar munt als ook *B* het doet (tweede kolom). Dus heeft *A* geen dominerende gedragslijn, geen beste gedragslijn onder alle omstandigheden. Hetzelfde geldt voor *B*, zoals men gemakkelijk kan nagaan.

Nu een iets ingewikkelder voorbeeld. We gaan naar de steegjes van Genua, waar we verrast worden door twee drukgebarende Italiaantjes *A* en *B* die elkaar toeschreeuwen en geregeld

knikkers heen en weer schuiven. Ze knikkeren niet, ze spelen *morra*. Het spel komt op het volgende neer. *A* en *B* tonen tegelijkertijd elk naar keuze 1 of 2 vingers en bovendien zèggen ze tegelijkertijd 1 of 2 als voorspelling van het aantal vingers dat de tegenstander gaat tonen. *A* heeft dus vier mogelijkheden:

1 vinger tonen en 1 zeggen
1 vinger tonen en 2 zeggen
2 vingers tonen en 1 zeggen
2 vingers tonen en 2 zeggen.

B heeft dezelfde vier mogelijkheden; er zijn dus $4 \times 4 = 16$ denkbare combinaties in totaal. Het betaalrooster bepaalt, dat als beide spelers correct het aantal opgestoken vingers van de tegenstander voorspellen, niemand iets krijgt. Het spel is dan remise. Hetzelfde geldt voor het geval waarin beide spelers het mis hebben. Maar wanneer een van de spelers het goed doet en de ander niet, dan ontvangt hij van zijn minder succesvolle tegenstander een aantal knikkers gelijk aan het totaal aantal opgestoken vingers, dus zijn eigen vingers plus die van zijn tegenstander.

Het is dus, al met al, een tamelijk ingewikkeld spel, waarin de Italianen echter een indrukwekkende snelheid weten te bereiken. Het is ook niet vrij van emoties. (De verleiding is al gauw aanwezig om de tegenstander ervan te beschuldigen te laat 1 of 2 te zeggen, dus pas te 'voorspellen' als het aantal vingers al is getoond, maar de gevolgen hiervan vallen buiten het bestek van dit boek.) Het betaalrooster is overigens niet bijzonder ingewikkeld. Het aantal door *A* te ontvangen knikkers blijkt uit het volgende overzicht:

Gedragslijn van *A*	Gedragslijn van *B*			
	zegt 1 toont 1	zegt 1 toont 2	zegt 2 toont 1	zegt 2 toont 2
zegt 1, toont 1	0	—3	2	0*
zegt 1, toont 2	3	0*	0	—4
zegt 2, toont 1	—2	0	0*	3
zegt 2, toont 2	0*	4	—3	0

De nullen zonder sterretje hebben betrekking op de gevallen, waarin beide partijen correct voorspellen. Bijv. de 0 linksboven correspondeert met de situatie, waarin beiden 1 vinger tonen

en 1 zeggen, dus correct voorspellen, dus geen knikkers uitwisselen. De 0 in de derde rij en de tweede kolom: *A* zegt 2 en dat wordt ook door *B* getoond, *B* zegt 1 en dat wordt ook door *A* getoond. De nullen met sterretje duiden op de situatie waarin beiden er naast zijn. Bijv. 0* rechtsboven: *A* zegt 1 maar *B* toont 2, *B* zegt 2 maar *A* toont 1. In alle overige gevallen worden knikkers van eigenaar verwisseld. Bijv. —3 in de eerste rij en de tweede kolom: *A* zegt 1 maar *B* toont 2, *A* is er dus naast; *B* zegt 1 en *A* toont 1, *B* heeft het dus bij het rechte eind. Volgens de regels van het spel ontvangt *B* dan evenveel knikkers als er vingers opgestoken zijn, dus $2 + 1 = 3$, vandaar de —3 voor de door *A* te ontvangen knikkers in het rooster.

Ook hier ziet men gemakkelijk, dat er geen dominerende gedragslijn is. Kiest *B* bijv. de gedragslijn behorend bij de eerste kolom (zegt 1, toont 1), dan is het beste wat *A* kan doen 'zegt 1, toont 2', want daarmee wint hij 3 knikkers terwijl hij anders niets zou winnen of 2 zou verliezen. Maar diezelfde handelwijze zou *A* tot het verlies van 4 knikkers veroordelen als *B* de gedragslijn van de vierde kolom zou volgen, terwijl er in die situatie voor *A* 3 knikkers te winnen waren. Ook *B* heeft geen dominerende gedragslijn. Dat volgt onmiddellijk uit het feit, dat *A* en *B* in dit spel in een volkomen symmetrische situatie verkeren. Wat waar is voor de een moet dus ook waar zijn voor de ander.

Tot zover de knikkers en de zilveren guldens. We hebben dus gezien dat dominerende gedragslijnen afwezig kunnen zijn, maar bovendien zijn onze luchthartige voorbeelden geschikt om een aantal begrippen te introduceren. In de eerste plaats het *spel* zelf. Dat is niets anders dan het geheel van de regels die in acht moeten worden genomen: dus uit welke mogelijkheden de spelers hun keus moeten maken en hoe na afloop de buit verdeeld wordt. Een spel bestaat steeds uit een of meer successieve *zetten* voor elk van de spelers. Bij kruis of munt heeft elke speler maar één zet: kruis boven of munt boven. Dat geldt ook voor morra. Weliswaar doet elke speler twee dingen (hij zegt iets en hij toont iets), maar dat doet hij tegelijkertijd, zodat er in feite maar één zet per speler is. Er zijn echter talrijke spelen met meer zetten. Schaken is een voorbeeld. Een eenvoudiger voorbeeld is *boter-kaas-en-eieren* en het loont de moeite daar enige aandacht aan te schenken. Er zijn weer twee spelers, *A*

en *B*, die beide turen naar een veld dat in 9 hokjes is verdeeld:

1	2	3
4	5	6
7	8	9

A begint en bezet een van de hokjes, bijv. 1. Dat is de eerste zet van *A*. Vervolgens bezet *B* een ander veld, bijv. 3; dat is zijn eerste zet. Daarna komt *A* weer aan de beurt. Hij bezet zegge 9, dat is dan zijn tweede zet. Enzovoorts. Het gaat er voor elk van de spelers om het eerst een complete rij te bezetten (bijv. 4-5-6) of een complete kolom (bijv. 1-4-7) of een diagonaal (1-5-9 of 3-5-7); en uiteraard om te verhinderen dat de ander dat het eerst doet. Men ziet gemakkelijk, dat de speler die begint 5 zetten voor de boeg heeft en de ander 4; er zijn immers 9 hokjes te bezetten.

Wanneer er verscheidene successieve zetten zijn, dan brengt dat een complicatie met zich mee, die echter aan de hand van wat we sinds het vorige hoofdstuk weten gemakkelijk valt op te lossen. In beginsel althans. In de loop van het spel krijgt elke speler immers nieuwe en relevante informatie, want hij ziet wat zijn tegenstander bezig is te doen. Kennelijk verdient het aanbeveling van meet af aan rekening te houden met het feit, dat die informatie er in de toekomst zal zijn, hetgeen impliceert dat we met *strategieën* dienen te werken. Bij boter-kaas-en-eieren zou *A* bijv. de volgende strategie kunnen kiezen:

(1) Als eerste zet bezet ik hokje 5
(2a) Als *B* dan 1 of 8 bezet, bezet ik 9
(2b) Als *B* dan 3 of 4 bezet, bezet ik 7
(2c) Als *B* dan 6 of 7 bezet, bezet ik 3
(2d) Als *B* dan 2 of 9 bezet, bezet ik 1
(3a) Als (2a) zich voordoet en als *B* in zijn tweede zet 4 bezet dan . . .

Dat zal in de practijk rijkelijk ingewikkeld worden, maar daar gaat het ons hier niet om. Begripsmatig is het eenvoudig. En het heeft een aanzienlijke vereenvoudiging van ons probleem tot gevolg: we hebben het spel van verscheidene successieve zetten per speler in feite teruggebracht tot een spel met één enkele zet per speler! Dat hebben we nl. al in het vorige hoofdstuk bij het schaken gezien (blz. 187). Als beide spelers een

strategie kiezen ligt daarmee het verloop van het spel vast. Die strategieën dienen aan het begin van het spel gekozen te worden. Het gaat voor elk van de twee spelers dan om één enkele keuze, nl. die van de te volgen strategie; en die ene keuze komt in feite neer op de enkele zet als het spel maar één zet per speler kent. Wij zullen daarom in hetgeen volgt de spelers uit strategieën laten kiezen, zodat de theorie dan toepasbaar is op spelen met een willekeurig aantal zetten.

3. MINIMAX EN ZADELPUNT

We nemen een heel eenvoudig voorbeeld: twee spelers, A en B; A heeft te kiezen uit twee strategieën, A_1 en A_2, en B ook uit twee, B_1 en B_2. Wat A van B in de vier mogelijke situaties ontvangt (in guldens) blijkt uit het volgende overzicht:

Strategieën	Strategieën van B	
van A	B_1	B_2
A_1	9	2
A_2	4	3

A krijgt dus onder alle omstandigheden een zeker bedrag van B, minimaal 2 en maximaal 9 gulden; hij is erin geïnteresseerd dat bedrag zo groot mogelijk te doen zijn. Voor B geldt het omgekeerde: hij moet altijd betalen, maar liefst (vanuit zijn standpunt) zo min mogelijk. Het spel is dus een beetje eenzijdig in het nadeel van B, maar dit doet er voor ons doel niet toe.

Hoe lossen we dit op? Laten we bij A beginnen. Hij is gefascineerd door die 9 gulden en om die te krijgen moet hij strategie A_1 kiezen. Hij slaagt in die opzet wanneer B_1 de keus van B is. Maar dat heeft A niet in de hand! Mocht B er achter komen dat A's keus op A_1 is gevallen (door spionnage, telepathie of wat ook), dan zal hij onmiddellijk B_2 kiezen; dan hoeft hij nog maar 2 gulden af te dragen. Maar is B's keus eenmaal op B_2 gevallen, dan wordt het voor A voordeliger A_1 te laten schieten en het op A_2 te houden – daardoor stijgt zijn opbrengst weer van 2 naar 3 gulden. En dan hebben we een merkwaardige situatie bereikt: noch A noch B hebben aanleiding hun positie te wijzigen. A niet omdat gegeven B_2 hij met A_2 een hogere opbrengst krijgt dan met A_1. B niet omdat gegeven A_2 hij

met B_2 minder te betalen heeft dan met B_1. Kennelijk hebben we rechtsonder in het betaalrooster een soort evenwichtspunt bereikt.

Voor een goed begrip diene, dat het geen Wet van Meden en Perzen is dat deze oplossing voor beide spelers altijd de optimale is. Wanneer A ervan zou kunnen uitgaan, dat B om wat voor reden dan ook aan de strategie B_1 verknocht is, dan doet hij er verstandig aan A_1 te kiezen en aldus een winst van 9 gulden te maken. (Het zou bijzonder dom van B zijn, want B_1 is voor hem onder alle omstandigheden slechter dan B_2.) En wanneer B ervan kan uitgaan, dat A aan niets anders dan die 9 gulden denkt en dus alleen aan A_1, dan behoeft hij slechts 2 gulden te betalen, niet 3. Maar deze redeneringen worden door de gangbare speltheorie geschuwd. Die theorie is zonder meer argwanend en gaat ervan uit, dat de spelers bang zijn voor het uitlekken van hun keus. Dat bleek al uit de redenering van de vorige alinea; en het is ook niet onredelijk. Hoezeer spelletjes ook in vriendschap gespeeld mogen worden, in het spel zèlf zijn de spelers volstrekte vijanden. Wat de een wint verliest de ander!

Het is nuttig het spel van deze gezichtshoek uit nog eens te bekijken. A redeneert dan: als ik de strategie A_1 kies, loop ik het risico maar 2 gulden te krijgen, dus het *minimum* van de eerste *rij* van het betaalrooster. En, zo vervolgt hij, door A_2 te kiezen loop ik het risico om niet meer dan het minimum van de tweede rij te krijgen, 3 gulden. B van zijn kant redeneert, dat hij door B_1 te kiezen het risico loopt 9 gulden te moeten betalen, dus het *maximum* van de eerste *kolom*. En door B_2 te kiezen loopt hij het risico van een verlies van 3 gulden, het maximum van de tweede kolom. Laten we deze rijminima en kolommaxima in de marges van het betaalrooster aangeven:

Strategieën van A	Strategieën van B: B_1	B_2	Rij-minimum
A_1	9	2	2
A_2	4	3	3*
Kolommaximum	9	3*	

Waar A naar streeft (onder onze argwanende veronderstelling)

is het *zo hoog mogelijk* doen zijn van de opbrengst onder de *ongunstigste* omstandigheid, dus wanneer zijn strategiekeuze aan *B* bekend wordt. Hij zal dus de *rijminima maximeren*, vandaar dat we van de *minimax-oplossing* plegen te spreken, die in dit geval naar A_2 voert met als minimale opbrengst 3; deze 3 is in de rechterkolom met een ster aangegeven. *B* van zijn kant doet precies hetzelfde. Aangezien dit betaalrooster zijn verliezen aangeeft en zijn strategieën kolomsgewijs specificeert, zal hij de kolom met het kleinste maximum kiezen, dus de strategie B_2. Dit leidt tot een maximaal te betalen bedrag van 3 gulden (in de onderste rij met een ster aangegeven), dus precies hetzelfde als hetgeen *A* minimaal te winnen heeft. Gemakkelijk ziet men, dat de twee gesterde drieën corresponderen met dezelfde 3 binnen de omlijning. Dat is de oplossing volgens het minimax-criterium. Door zich aan dat criterium te houden verzekert *A* dat hij 3 gulden ontvangt; het kàn ook meer zijn, dus voordeliger voor hem, nl. wanneer *B* zich niet aan minimax houdt (*A* krijgt dan 4 gulden). Door te handelen volgens het minimax-criterium beperkt *B* zijn verlies tot 3 gulden. Zijn verlies kan kleiner zijn, nl. wanneer *A* zich niet aan minimax houdt (*B* hoeft dan maar 2 gulden te betalen). Het bedrag dat *A* van *B* ontvangt als beiden volgens minimax te werk gaan staat bekend als de *waarde* van het spel; dus hier 3 gulden.[1]

De minimax-oplossing, dus de 3 rechtsonder in het omlijnde stuk van het betaalrooster, staat bekend als het *zadelpunt* van het rooster; d.w.z., het is een minimum wanneer we naar de rij kijken en een maximum wanneer we naar de kolom kijken. De betekenis van die term wordt duidelijker wanneer we *A* en *B* elk nog een derde strategie beschikbaar stellen:

1. Merk op dat we bij het wedden op Duindigt in Hoofdstuk 1 ook volgens minimax te werk zijn gegaan. Het criterium was immers: verdeel de beschikbare 57 gulden op een dusdanige manier, dat de winst in de ongunstigste omstandigheid nog zo groot mogelijk is. Toch is er een belangrijk verschil als gevolg van het feit, dat er op Duindigt geen echte tegenstander is van het type zoals in dit hoofdstuk beschouwd. Hier is er een volstrekte tegenstander: wat U wint verliest hij. Er is dan alle reden om argwanend te zijn. Op Duindigt niet; die paarden hebben niets tegen U, evenmin het lot dat hun succes bepaalt. Vandaar dat het minimax-criterium op Duindigt toegepast eerder pessimistisch dan argwanend is.

Strategieën van A	Strategieën van B B_1	B_2	B_3	Rij-minimum
A_1	9	2	1	1
A_2	4	3	5	3*
A_3	0	1	7	0
Kolommaximum	9	3*	7	

Het hoogste rijminimum blijft 3 en het laagste kolommaximum blijft ook 3, evenals in het vorige geval corresponderend met resp. A_2 en B_2, zodat de minimax-oplossing niet verandert. Bezie nu A's opbrengst in geval hij A_2 kiest. De drie opbrengsten zijn (in de volgorde B_1, B_2, B_3): 4, 3, 5. Hij daalt dus eerst van 4 naar 3 en stijgt dan van 3 naar 5. Neem dan A's opbrengst in geval B_2 de keus van B is. Weer drie opbrengsten, nl. 2, 3, 1. Kennelijk *stijgt* de opbrengst dan eerst van 2 naar 3 (dezelfde 3 van zojuist), daarna daalt hij naar 1. Het getal 3 middenin het betaalrooster heeft dus een ietwat asymmetrische positie: het vormt een dal wanneer we de tweede rij bekijken, een top wanneer we de tweede kolom bekijken. Maar dat is precies als bij een zadel (en vandaar de term zadelpunt). Leggen we het zadel op een paard, dan is het midden van het zadel een dal wanneer we ons oog de lengte-as van het paard laten volgen, maar een top wanneer we dwars daarop kijken. Hetzelfde hebben we bij een bergpas. Volgen we de weg, dan klimmen we tot we op de pas zijn, daarna dalen we; volgens de wegrichting is een pas dus een top. Maar zijn we op de pas en kijken we naar links en naar rechts, dan zien we aan beide zijden een berg; in die richting bezien is een pas dus een dal.

Hoewel we het voorbeeld van de luchtvaartmaatschappijen van het begin van dit hoofdstuk daar al tot oplossing hebben gebracht, is het aardig om nog even te laten zien, dat die oplossing een minimax-oplossing is. We spreken weer gemakshalve over strategieën (ook al is er maar één zet per maatschappij). Het betaalrooster van blz. 204 is opnieuw afgedrukt op blz. 214 (bovenaan) maar de rijminima en de kolommaxima zijn er nu aan toegevoegd (resp. in de rechterkolom en de onderste rij). U ziet, het hoogste rijminimum correspondeert weer met het laagste kolommaximum, we komen rechtsonder uit. Beide partijen gaan adverteren – hetgeen we al wisten. De waarde van het spel is 50 passagiers per jaar.

Strategieën van A	Strategieën van B 1 (niets)	2 (films)	3 (advert.)	Rij-minimum
1 (niets)	0	—100	—300	0
2 (films)	100	0	—100	—100
3 (adverteren)	350	125	50	50*
Kolommaximum	350	125	50*	

4. GEMENGDE STRATEGIEËN

Wellicht heeft de hier gepropageerde argwaan aanstekelijk gewerkt en gaat de lezer als volgt redeneren. In dat voorbeeldje van twee strategieën voor *A* en *B* elk hoef ik helemaal niet zo ingewikkeld te werk te gaan om tot de oplossing te komen, of we hem nu minimax noemen of iets anders. Want B_1 is een strategie die, wat *A* ook doet, voor *B* altijd slechter is dan B_2. *A* weet even goed dat B_1 door B_2 gedomineerd wordt, dat kan hij uit het betaalrooster aflezen; en als hij dat eenmaal weet komt hij automatisch tot A_2. Volmaakt correct, maar we moeten één kanttekening maken: het argument gaat niet meer op wanneer de strategieën A_3 en B_3 tot de mogelijkheden gaan behoren. B_1 wordt dan niet langer door B_2 gedomineerd, omdat B_1 voordeliger is wanneer A_3 de keus van *A* is. En *B* kan niet zonder meer uitsluiten dat *A*'s keus op A_3 valt, omdat die strategie vanuit *A*'s standpunt het voordeligst is wanneer B_3 de keus van *B* is.

Maar het is waar dat er zorgen zijn, zij het van andere aard. Laten we die illustreren aan de hand van het simpele kruis-of-munt spel:

Strategieën van A	Strategieën van B kruis	munt	Rij-minimum
kruis	1	—1	—1
munt	—1	1	—1
Kolommaximum	1	1	

Wat *A* minimaal te verwachten heeft is steeds —1, of hij nu kruis toont of munt. Wat moet hij nu doen? Voor *B* is het probleem precies hetzelfde: wat hij ook doet, hij moet in de ongunstigste situatie altijd 1 gulden betalen. Om hier wat klaar-

heid te scheppen zullen we te werk gaan op dezelfde manier als waarop we indertijd tot de uitkomst 3 kwamen. Neem aan dat *A* overweegt kruis te tonen. Doet B het ook, dan ontvangt *A* een gulden; maar mocht *B* achter *A*'s plan komen, dan zal hij zeker munt tonen met als gevolg dat *A* een gulden moet betalen. Onder die omstandigheid wordt het echter voor *A* aantrekkelijk om ook maar munt te tonen, dan krijgt hij de gulden. Maar als *B* dààr achter komt zal hij kruis tonen; en dan wil *A* ook weer kruis tonen. U ziet het, het circuit is rond, we zijn in de uitgangssituatie teruggekeerd.

Kennelijk is wat voor *A* voordelig is nu *volledig* afhankelijk van wat *B* gaat doen. Zou *A* er op de een of andere manier achter kunnen komen dat *B* een zekere voorkeur heeft voor bijv. kruis, dan wordt het de moeite waard daarmee rekening te houden. Neem bijv. aan, dat het spel een groot aantal malen gespeeld wordt en dat *B* in gemiddeld 3 op de 4 gevallen kruis toont. Dat kost hem geld als *A* daar achter komt! Want als *A* dan systematisch kruis gaat tonen, zal hij het spel winnen in 3 op de 4 gevallen, verliezen in maar 1 op de 4 gevallen, en dus gemiddeld genomen een winst van een halve gulden per spel maken. Het is blijkbaar bijzonder gevaarlijk van enige voorkeur voor kruis of voor munt blijk te geven. En dat geldt zowel voor *A* als voor *B*, want beide spelers verkeren in dezelfde positie. Bovendien geldt, dat ook regelmaat gevaarlijk is. Stel bijv., dat *B* er een gewoonte van maakt om de andere keer kruis boven te leggen. Dan is er geen sprake van voorkeur, maar toch ook daar kan *A* achter komen. Dus: geen voorkeur en geen regelmaat.

Hoe voorkomt een speler dat zijn tegenstander er achter komt wat hij gaat doen? Door het, tot op het laatste moment, zèlf niet te weten! Stel bijv., dat *A* er een gewoonte van maakt zijn gulden onder tafel op te gooien en vervolgens op tafel te leggen. Daarmee verzekert hij, dat de kruisstrategie in de helft van de gevallen wordt gevoerd en de muntstrategie eveneens; wat het zal zijn is voor de tegenstander volmaakt onvoorspelbaar (en voor hemzelf ook). Dit is wat in de speltheorie bekend staat als een *gemengde strategie*. Wat *A* zal doen bij ieder concreet spel staat niet van tevoren vast. Het kan kruis zijn, het kan munt zijn; dat is niet met zekerheid te zeggen, ook niet door *A* zelf. Het enige dat wel met zekerheid gezegd kan worden is dat beide alternatieven een kans $^1/_2$ hebben. Dus een

zeker 'mengsel' van de afzonderlijke strategieën, vandaar de term. Hoe staat het met de opbrengst van een dergelijke gemengde strategie? Dat is eenvoudig genoeg. Als *B* besluit kruis te tonen, levert de gemengde strategie een opbrengst 1 voor *A* met een kans $^1/_2$ (nl. wanneer die gemengde strategie in concreto op kruis blijkt neer te komen) en een opbrengst —1 (als hij op munt terecht komt), ook met een kans $^1/_2$. De verwachte opbrengst is dus $^1/_2 \times 1 + {^1/_2} \times (-1) = 0$ als *B* kruis toont. En als *B* munt toont heeft *A* ook een opbrengst 1 met kans $^1/_2$ en —1 met kans $^1/_2$, dus ook weer een verwachte opbrengst van nul. Kort gezegd, een gemengde strategie heeft geen bepaalde opbrengst met zekerheid; wij waarderen hem tegen de verwachte opbrengst.

Het gaat dus nu om de vraag wèlke gemengde strategie een speler dient te kiezen. In het bovenstaande voorbeeld is dat duidelijk genoeg. Beide spelers dienen een gemengde strategie met kans $^1/_2$ op kruis en kans $^1/_2$ op munt te volgen; dat is immers de enige manier om te vermijden dat de tegenstander kan profiteren van getoonde voorkeur. Maar het is niet altijd zo eenvoudig. Laten we iets aan het betaalrooster veranderen:

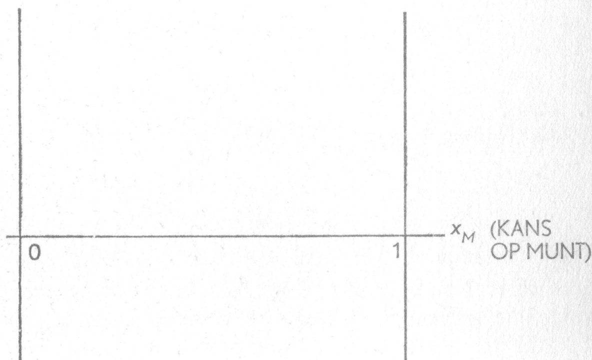
Strategieën van *A*	Strategieën van *B* kruis	 munt	Rij- minimum
kruis	2	—1	—1
munt	—1	1	—1
Kolommaximum	2	1	

De enige verandering is een verhoging van *A*'s winst van 1 tot 2 gulden in geval beide spelers kruis tonen. Het spel is nu niet langer symmetrisch; er zit voor *A* meer winst in dan voor *B*. De vraag is echter hoe *A* die winst waar kan maken. Het ligt voor de hand dat hij voorkeur voor kruis zal hebben, want alleen dan is er voor hem een mogelijkheid die 2 gulden te incasseren. Maar dat weet *B* ook en zijn voorkeur gaat dus uit naar munt, waardoor de winst van *A* zou dalen van 2 tot —1. En dàt weet *A* even goed! Wij dreigen dus weer in een cirkelsituatie te geraken en het ligt in de rede, dat we het weer bij gemengde strategieën moeten zoeken. Maar welke? Kennelijk moeten we de waarschijnlijkheden van kruis en van munt dan op een zo gunstig mogelijke manier vaststellen. En dat is dus nu de es-

sentie van het keuzeprobleem geworden: eerst ging het om de selectie van één van de beschikbare strategieën, nu gaat het om de specificatie van de waarschijnlijkheden van die afzonderlijke strategieën in de gemengde strategie. Wij zullen aantonen, dat het beste dat *A* nu kan doen (gemeten aan de verwachte opbrengst) is: met kans 0,4 kruis en met kans 0,6 munt. Dus minder vaak kruis, terwijl daar zijn hart naar uitging! De verwachte opbrengst zal 0,2 gulden blijken, dus 20 cent.

5. EEN GRAFISCHE OPLOSSING

Ons spel is zo eenvoudig, dat de oplossing met een tekening kan worden afgeleid. Die tekening lijkt veel op die, welke we in het vorige hoofdstuk voor de autofabrikant hebben gebruikt, maar de interpretatie is anders. Langs de horizontale as meten we de waarschijnlijkheid van munt in *A*'s gemengde strategie. Langs de verticale as meten we *A*'s verwachte opbrengst.



Het is nu van belang te beseffen, dat wàt *B* ook doet, of hij een gemengde strategie volgt of niet, hij uiteindelijk met kruis of munt voor de dag komt. Dat is het enige belangrijke voor *A*. Hij zal dus zijn verwachte opbrengst moeten bepalen onder die twee alternatieve veronderstellingen; een van de twee zal zeker gerealiseerd worden. Laten we er eerst vanuit gaan, dat *B* kruis toont. Laten we voorts x_M schrijven voor de kans op munt

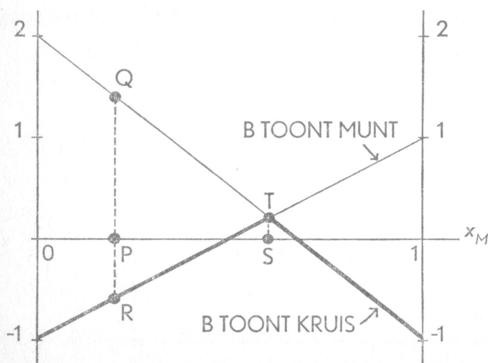
in A's gemengde strategie, dus $1 - x_M$ voor de kans op kruis. Is het inderdaad waar, dat B kruis toont, dan vinden we de volgende verwachte opbrengst van A aan de hand van de eerste kolom van het betaalrooster:

$$(1 - x_M)2 + x_M(-1) = 2 - 3x_M.$$

Neem nu aan, dat B munt toont. In dat geval geeft de tweede kolom van het betaalrooster te zien, dat A's verwachte opbrengst wordt:

$$(1 - x_M)(-1) + x_M \times 1 = -1 + 2x_M.$$

Dit zijn twee lineaire functies van x_M, die beide door rechte lijnen in onze figuur worden weergegeven. De eerste begint bij 2 (als $x_M = 0$) en eindigt bij -1 (als $x_M = 1$); de tweede begint bij -1 (als $x_M = 0$) en eindigt bij 1 (als $x_M = 1$). De figuur ziet er dus als volgt uit:



Wat dient A nu te doen? Welnu, zijn probleem is een zo goed mogelijke gemengde strategie te vinden, d.w.z. hij dient de waarde van x_M te vinden waarvoor de verwachte opbrengst *in het ongunstigste geval* nog zo groot mogelijk is. Bezie dan het punt P, corresponderend met een kans op munt (x_M) gelijk aan 0,2. Dan kan hij een verwachte opbrengst van PQ krijgen, nl. wanneer B kruis toont, maar die verwachte opbrengst kan ook negatief zijn ter grootte van PR, nl. wanneer B munt toont; en de laatste opbrengst is uiteraard de laagste. Het valt gemakkelijk in te zien, dat de beste procedure is in S te gaan zitten (corresponderend met het snijpunt T van de twee lijnen). Gaan

we immers nog verder naar rechts, dus verhogen we x_M nog meer, dan daalt de minimale verwachte opbrengst beneden het niveau *ST* i.v.m. de mogelijkheid dat *B* kruis gaat tonen.

Het gaat dus om het snijpunt van de twee lijnen, dat we vinden door de twee verwachte opbrengsten aan elkaar gelijk te stellen:

$$2 - 3x_M = -1 + 2x_M$$

of $5x_M = 3$ met als antwoord: $x_M = 0{,}6$. De kans op kruis moet dus slechts 0,4 zijn. En de verwachte opbrengst is dan

$$2 - 3 \times 0{,}6 = -1 + 2 \times 0{,}6 = 0{,}2,$$

dus 20 cent, zoals al eerder vermeld.

Voor *B* gaan we op volkomen analoge wijze te werk; we tekenen ook voor hem een dergelijke figuur en vinden dan, dat zijn beste strategie op het volgende mengsel neerkomt: kans 0,4 op kruis, kans 0,6 op munt. Toevallig dezelfde gemengde strategie dus. Die valt te verwezenlijken door een vijfkantig potlood te laten rollen. Er zijn aanmerkelijk meer geavanceerde methoden om keuzen op kansbasis te doen; daarvoor verwijzen we naar Hoofdstuk 10.

6. DE MINIMAX-STELLING

We gaan nu een iets groter voorbeeld bezien. *A* en *B* hebben 3 resp. 4 strategieën ter beschikking en het betaalrooster (in guldens uitgedrukt, door *B* aan *A* te betalen) ziet er als volgt uit:

Strategieën van *A*	Strategieën van *B*				Rij-minimum
	B_1	B_2	B_3	B_4	
A_1	10	1	7	8	1
A_2	3	4	5	9	3*
A_3	6	7	4	2	2
Kolommaximum	10	7*	7*	9	

Het hoogste rijminimum is 3, het laagste kolommaximum is 7 (met sterren aangegeven), dus verschillend. Er is geen zadelpunt. Dan is het altijd fout om het op een en dezelfde strategie te houden, zoals met de volgende 'cirkel' redenering eenvoudig kan worden geïllustreerd. Stel *A* houdt het op A_2, omdat hij daardoor een winst van 3 gulden voor zich garandeert. Dan is

B_1 het beste dat B kan doen; wat B moet betalen blijft dan beperkt tot die 3 gulden. Maar A dient dit te begrijpen en te concluderen dat, als inderdaad B_1 gespeeld wordt door B, het voor hem veel voordeliger wordt A_1 te spelen dan A_2; hij wint dan 10 gulden. En B begrijpt ook dat weer en mikt het op B_2: 1 gulden. Dan gaat A naar A_3: 7 gulden. Dan B naar B_4: 2 gulden. Dan A naar A_2: 9 gulden. En daarmee zijn we op het uitgangspunt teruggekeerd en gaat het van voren af aan. Deze situatie kan niet voorkomen als er een zadelpunt is, want dan raken we ergens in een evenwichtspositie. Maar meestal is er geen zadelpunt en moeten we het zoeken bij de gemengde strategieën.

Voor A gaat het erom een geschikt mengsel van zijn strategieën A_1, A_2, A_3 te vinden, dus om de specificatie van de drie kansen x_1, x_2, x_3 waarmee hij die strategieën in feite gaat voeren. Uiteraard moeten die kansen positief of nul zijn (symbolisch: $\geqq 0$) en hun som moet 1 zijn; A moet tenslotte enige strategie kiezen, welke dan ook. Dus:

(1) $x_1 \geqq 0, \quad x_2 \geqq 0, \quad x_3 \geqq 0$
$x_1 + x_2 + x_3 = 1.$

Dan moet hij de x-en met inachtneming van deze restricties zodanig kiezen, dat de minimale verwachte winst zo groot mogelijk is. We bezien eerst de verwachte winst onder de veronderstelling dat de B_1 de keus van B blijkt te zijn. Dan maakt A een winst van 10 met kans x_1, van 3 met kans x_2, en 6 met kans x_3; zie de eerste kolom van het betaalrooster. A's verwachte winst in geval B besluit de strategie B_1 te kiezen is dus:

$10x_1 + 3x_2 + 6x_3.$

Zo kunnen we doorgaan met de tweede, derde en vierde kolom van het betaalrooster, corresponderend met de strategiekeuzen B_2, B_3, B_4 van B. Het resultaat laat zich als volgt samenvatten:

Strategiekeuze van B	Verwachte winst van A
B_1	$10x_1 + 3x_2 + 6x_3$
B_2	$x_1 + 4x_2 + 7x_3$
B_3	$7x_1 + 5x_2 + 4x_3$
B_4	$8x_1 + 9x_2 + 2x_3$

Dus vier verwachte winsten, elk behorend bij een van de keuzen die B kan blijken te doen. Het gaat erom de laagste van de vier nog zo groot mogelijk te doen zijn. Laten we dan w schrijven

voor de laagste verwachte winst, dus voor het *minimum* van dit viertal. Dan volgt onmiddellijk uit de definitie van het minimum:

$$(2)\quad \begin{aligned} 10x_1 + 3x_2 + 6x_3 &\geqq w \\ x_1 + 4x_2 + 7x_3 &\geqq w \\ 7x_1 + 5x_2 + 4x_3 &\geqq w \\ 8x_1 + 9x_2 + 2x_3 &\geqq w. \end{aligned}$$

Het probleem van A is dus: kies x_1, x_2, x_3 zodanig dat daardoor w (de laagste verwachte winst) zo groot mogelijk is, echter met inachtneming van de vier restricties (1) en de vier restricties (2).

We zullen dit probleem niet direct oplossen; dat zou niet fair zijn tegenover B, die per slot van rekening met hetzelfde probleem worstelt. Hij heeft vier strategieën ter beschikking, B_1, B_2, B_3, B_4, en dus is het zijn taak vier waarschijnlijkheden voor zijn gemengde strategie te specificeren: y_1, y_2, y_3, y_4. Uiteraard moeten ook die niet-negatief zijn en som 1 hebben:

$$(3)\quad \begin{aligned} &y_1 \geqq 0, \quad y_2 \geqq 0, \quad y_3 \geqq 0, \quad y_4 \geqq 0 \\ &y_1 + y_2 + y_3 + y_4 = 1. \end{aligned}$$

Ook B is erop uit zijn laagste verwachte winst zo groot mogelijk te doen zijn. Of ook (aangezien het betaalrooster hem weinig illusies kan laten): zijn hoogste verwachte verlies zo klein mogelijk. We vinden gemakkelijk, dat als A_1 de keus van A is, dit verwachte verlies gelijk is aan

$$10y_1 + y_2 + 7y_3 + 8y_4,$$

zie de eerste rij van het betaalrooster. Voor de tweede en derde rij (corresponderend met A_2 en A_3 als strategiekeuzen van A) gaat het natuurlijk analoog. Schrijf dan v voor het *maximale* verwachte verlies van B; dan volgt uit de definitie van het maximum:

$$(4)\quad \begin{aligned} 10y_1 + y_2 + 7y_3 + 8y_4 &\leqq v \\ 3y_1 + 4y_2 + 5y_3 + 9y_4 &\leqq v \\ 6y_1 + 7y_2 + 4y_3 + 2y_4 &\leqq v. \end{aligned}$$

Het probleem voor B is dan: kies y_1, y_2, y_3, y_4 zodanig dat daardoor v (het grootste verwachte verlies) zo klein mogelijk is, echter met inachtneming van de vijf restricties (3) en de drie restricties (4).

Dat zijn dus twee verschillende problemen, één voor A en één voor B; voor personen dus, die weinig reden hebben om met elkaar in harmonie te leven, want wat de een wint dat verliest de ander. Het merkwaardige is nu, dat er toch een soort van harmonie bestaat, zij het in beperkte zin – die van het

evenwicht in een zadelpunt. Echter geen zadelpunt in de sfeer van afzonderlijke strategieën, maar in die van gemengde strategieën. Dit is de minimax-stelling van Von Neumann, die geldig is voor alle nulspelen van twee personen, hoeveel strategieën er ook zijn en hoe het betaalrooster er ook uitziet. Deze stelling komt op het volgende neer. Er is voor *A* een bepaalde gemengde strategie (dus zekere waarschijnlijkheden x_1, x_2, ...) en ook voor *B* is er een bepaalde gemengde strategie (dus zekere kansen y_1, y_2, ...) zodanig dat

(i) *A* door het spelen van de hem aangewezen strategie er zeker van kan zijn dat de verwachte winst minimaal *w* is en, evenzo, *B* door het spelen van de zijne er zeker van kan zijn dat het verwachte verlies maximaal *v* is;

(ii) de minimale verwachte winst *w* van *A* is gelijk aan het maximale verwachte verlies *v* van *B*;

(iii) als *A* de hem aangewezen strategie speelt maar *B* niet, dan kàn de verwachte winst van *A* boven het minimum *w* uitkomen; en evenzo, als *B* de zijne speelt maar *A* niet, dan kàn het verwachte verlies van *B* beneden het maximum *v* liggen;

(iv) het minimum *w* en het maximum *v* [blijkens (ii) zijn ze aan elkaar gelijk] liggen steeds ergens tussen het maximum van de rijminima van het betaalrooster en het minimum van de kolommaxima.

Laten we dan zien, wat dit voor ons voorbeeld betekent. De oplossing van *A*'s gemengde strategie is:[1] *voor noot zie p. 223*

$$x_1 = \frac{10}{42}, \quad x_2 = \frac{9}{42}, \quad x_3 = \frac{23}{42}.$$

Alle *x*-en zijn positief en hun som is inderdaad 1. Op blz. 220 is de verwachte winst van *A* voor willekeurige *x*-en aangegeven, wanneer *B* zijn vier verschillende keuzen maakt. Vullen we de numerieke waarden in, dan vinden we achtereenvolgens:

$$B_1: \quad 10 \times \frac{10}{42} + 3 \times \frac{9}{42} + 6 \times \frac{23}{42} = \frac{265}{42}$$

$$B_2: \quad 1 \times \frac{10}{42} + 4 \times \frac{9}{42} + 7 \times \frac{23}{42} = \frac{207}{42}$$

$$B_3: \quad 7 \times \frac{10}{42} + 5 \times \frac{9}{42} + 4 \times \frac{23}{42} = \frac{207}{42}$$

$$B_4: \quad 8 \times \frac{10}{42} + 9 \times \frac{9}{42} + 2 \times \frac{23}{42} = \frac{207}{42}$$

Het resultaat toont in één oogopslag, dat de verwachte winst van A minimaal 207 / 42 is (dus bijna 5 gulden). Voor B is de oplossing:

$$y_1 = 0, \quad y_2 = \frac{15}{42}, \quad y_3 = \frac{24}{42}, \quad y_4 = \frac{3}{42}.$$

De waarde nul voor y_1 betekent dat de strategie B_1 niet gespeeld wordt. (We zullen dadelijk zien dat B daar goede redenen voor heeft.) Het verwachte verlies van B voor de verschillende keuzen van A wordt dan:

$$A_1: \quad 10 \times 0 + 1 \times \frac{15}{42} + 7 \times \frac{24}{42} + 8 \times \frac{3}{42} = \frac{207}{42}$$

$$A_2: \quad 3 \times 0 + 4 \times \frac{15}{42} + 5 \times \frac{24}{42} + 9 \times \frac{3}{42} = \frac{207}{42}$$

$$A_3: \quad 6 \times 0 + 7 \times \frac{15}{42} + 4 \times \frac{24}{42} + 2 \times \frac{3}{42} = \frac{207}{42}$$

Het verwachte verlies van B is dus maximaal 207 / 42. Hiermee hebben we de punten (i) en (ii) van de minimax-stelling geïllustreerd: A maakt minimaal een verwachte winst van w = 207 / 42 en B lijdt maximaal een verwacht verlies van v =207 / 42, dus $v = w$. Ook zien we, dat punt (iii) opgaat: als B zo dom zou zijn systematisch B_1 te gaan spelen, dan stijgt de verwachte winst van A tot 265 / 42. De hem aangewezen strategie is nu net niet zo dom: dat is een gemengde strategie waarvan het mengsel B_1 helemaal niet bevat, want $y_1 = 0$. En ook aan punt (iv) is voldaan: het hoogste rijminimum van het betaalrooster is 3, het laagste kolommaximum is 7, en $v = w$ ligt hier tussenin. Deze verwachte opbrengst (w = 207 / 42) staat bekend als de *waarde* van het spel in geval het betaalrooster geen zadelpunt heeft.

Nog een enkel woord over het zadelpunt. Heeft het betaalrooster een dergelijk punt, dan zijn er steeds twee strategieën, één voor A en één voor B, die de minimax-oplossing vormen. Is er geen zadelpunt, dan wordt er in feite zo'n punt gemaakt met behulp van gemengde strategieën. De stelling van Von

1. De afleiding wordt hier niet gegeven. Wie Hoofdstuk 1 gelezen heeft en zich Duindigt herinnert, zal aanvoelen dat het op lineair programmeren neerkomt.

Neumann garandeert immers het bestaan van gemengde strategieën met precies hetzelfde effect: er bestaat nl. evenwicht in de zin dat geen van de spelers zijn positie verbeteren kan (gemeten aan de verwachte opbrengst) wanneer de ander zich aan de hem voorgeschreven gemengde strategie houdt. En dit is volledig analoog aan de zadelpuntoplossing voor de afzonderlijke strategieën.

7. OVER MEER SPELERS EN NIET-NULSPELEN

Tot slot een enkel voorbeeld van een spel buiten de sfeer van twee personen en som nul van de opbrengsten. Het is het probleem van een varkensfokker, de heer Nachtegaal. Zijn probleem is: hoeveel varkens moet ik per jaar fokken? Naast Nachtegaal worden nog een duizend andere fokkers geconfronteerd met dit zelfde probleem. Laten we nu voor het gemak aannemen, dat iedere fokker kan besluiten weinig (10) of véél (25) varkens te fokken. Als *alle* boeren weinig fokken, komen er vrij weinig varkens op de markt; deze varkens zullen dan stuk voor stuk goede prijzen maken, nl. 400 gulden per stuk. Dat komt neer op een winst van 150 gulden per stuk, omdat het fokken 250 gulden kost. De winst voor Nachtegaal is dan $10 \times 150 = 1500$ gulden. De 1000 andere boeren maken ook ieder een winst van 1500 gulden voor hun varkens. Stel nu, dat Nachtegaal zou besluiten, veel, dus 25, varkens te gaan fokken, terwijl alle andere boeren weinig blijven fokken. Die 15 varkens meer zullen de marktprijs voor varkens niet noemenswaard beïnvloeden, en Nachtegaal maakt dus een winst van $25 \times 150 = 3750$ gulden! Wat gebeurt er echter als *alle* fokkers, inclusief Nachtegaal, 25 varkens gaan fokken per jaar? Er komt dan een heel groot aanbod op de markt en de prijs daalt tot 275 gulden per varken, zodat er nog slechts een winst van 25 gulden per varken resteert. En een totale winst van $25 \times 25 = 625$ gulden voor ieder. En als, onder deze omstandigheden, Nachtegaal zou besluiten slechts 10 varkens te fokken, zou hij slechts een winst van $10 \times 25 = 250$ gulden beuren.

Nachtegaal beschikt dus over 2 mogelijkheden, nl. $K10$ en $K25$ (d.w.z. Nachtegaal fokt 10 resp. 25 varkens). Voor de anderen geldt precies hetzelfde: $A10$ en $A25$. Het betaalrooster,

gemeten in guldens winst op varkens per boer per jaar, ziet er dus als volgt uit:

	*A*10	*A*25
*K*10	Nachtegaal: 1500 gulden Anderen: 1500 gulden	Nachtegaal: 250 gulden Anderen: 625 gulden
*K*25	Nachtegaal: 3750 gulden Anderen: 1500 gulden	Nachtegaal: 625 gulden Anderen: 625 gulden

Uit dit overzichtelijke schema concludeert Nachtegaal direct, dat hij veel varkens moet produceren. Wàt de andere fokkers ook doen, door zelf 25 varkens te produceren is hij altijd beter uit. Hij wint immers 3750 i.p.v. 1500 gulden als de anderen weinig, en 625 i.p.v. 250 gulden als de anderen veel varkens fokken. Kortom, *K*25 domineert *K*10. Het probleem van Nachtegaal is nu opgelost: hij fokt 25 varkens per jaar (of, meer algemeen, zoveel hij kan).

Nu het drama. Ieder van de individuele varkensfokkers zal een dergelijk betaalrooster kunnen maken. En iedere fokker zal om precies dezelfde reden als Nachtegaal besluiten veel varkens te gaan fokken. En als ze allemaal veel varkens gaan fokken, dan maken zij allen 625 gulden winst. Terwijl als ze allemaal zouden hebben besloten weinig varkens te produceren ieder een winst van 1500 gulden zou maken! De beslissing die voor de individuele boer op zichzelf de juiste is, is voor alle boeren tezamen onjuist. Het is een der redenen waarom velen menen dat de overheid in de landbouw van tijd tot tijd regelend moet optreden. En, in de practijk, bijna overal inderdaad regelend optreedt.

Nu nog het een en ander over spelen voor meer dan twee personen. Een der verwikkelingen die hierbij kan optreden is dat het in de regel voordelig is voor een aantal spelers zich aaneen te sluiten. Zij vormen dan een soort kongsi of belangengemeenschap; wanneer zij dit gedaan hebben is een nieuw spel ontstaan met minder deelnemers. Trouwens, zelfs indien er slechts twee spelers zijn is dit soort kartelvorming niet uitgesloten. De luchtvaartondernemingen die op een bepaalde route concurreren kunnen best (en zullen in de regel) op die routes gaan samenwerken. Wanneer het een nulspel is kunnen beide spelers er niet tegelijkertijd beter op worden; maar veel spelen in de practijk zijn niet-nulspelen.

Het oorspronkelijke boek over speltheorie, meer dan 600 pagina's dichtbedrukte tekst, is dat van Von Neumann en Morgenstern [1], dat aan het einde van het vorige hoofdstuk al genoemd is. De boeken van McKinsey [2] en van Luce en Raiffa [3] zijn later geschreven, eenvoudiger, maar toch niet werkelijk eenvoudig voor de wiskundige leek. Het grappigste, en voor iedereen begrijpelijke boek, dat echter veel moeilijke problemen onbesproken laat (wat ook voor dit hoofdstuk geldt!), maar de gedachtengang duidelijk weergeeft, is dat van Williams [4].

[1] Neumann, J. von, and O. Morgenstern, *Theory of Games and Economic Behavior*. Third Edition. Princeton University Press, Princeton. 1953.

[2] McKinsey, J. C. C., *Introduction to the Theory of Games*. McGraw-Hill Book Company, Inc., New York. 1952.

[3] Luce, R. D., and H. Raiffa, *Games and Decisions*. John Wiley and Sons, Inc., New York. 1957.

[4] Williams, J. D., *The Compleat Strategyst*. McGraw-Hill Book Company, Inc., New York. 1954.

9. WACHTEN

1. HET PROBLEEM

Het probleem waarmee wij ons in dit hoofdstuk zullen bezig houden wordt goed geïllustreerd door het volgende voorbeeld. In een bepaald postkantoor is één loket dat zich bezig houdt met het behandelen van girozaken. Het loket gaat 's morgens om half negen open. Er zijn dan twee mogelijkheden: de eerste 'klant' meldt zich onmiddellijk of de loketbeambte begint zijn taak met niets te doen te hebben. In het laatste geval zal de eerste klant na kortere of langere tijd binnenkomen. Hij wordt dan onmiddellijk geholpen. Op het ogenblik dat hij gereed is doen zich t.a.v. de tweede klant drie verschillende mogelijkheden voor: hij is inmiddels binnengekomen en heeft moeten wachten, hij komt precies op het ogenblik van vertrek van de eerste klant binnen, of hij komt pas later zodat de loketbeambte weer enige tijd niets te doen heeft.

Wanneer de klant moet wachten is dat voor hem onplezierig. Hij had zijn tijd nuttiger of aangenamer kunnen besteden. Anderzijds, wanneer de loketbeambte niets te doen heeft, is dit voor de P.T.T. onaangenaam. Hij moet worden doorbetaald, zonder dat hij enige prestatie levert. (Dat is niet helemaal waar, hij doet ander werk zoals opruimen e.d.; maar het is duidelijk, dat hij vanuit het standpunt van de P.T.T. productiever bezig is als hij klanten helpt.) De ideale situatie verlangt dat iedere keer een klant binnenkomt als zijn voorganger net vertrekt; dat is ideaal voor beide partijen, want geen klant behoeft dan te wachten en de P.T.T. heeft geen leegloop. Maar dit ideaal is niet te verwezenlijken.

Het probleem dat men moet wachten is intussen veel algemener. Wij wachten bij de kapper, bij de bus- of tramhalte, voor de pont, op de toegangswegen van de Maastunnel. Machines wachten op reparatie, vliegtuigen wachten op het vrij komen van een landingsbaan, schepen op een loods of op laad- en losfaciliteiten, treinen voor onveilige signalen. In de meeste gevallen zou dit wachten voorkomen kunnen worden door uitbreiding van de bedieningsinstallatie. De kapper zou zijn zaak kunnen vergroten en extra bedienden aantrekken. Kruisingen met stoplichten kan men vervangen door ongelijk-

vloerse kruisingen, ponten kunnen vervangen worden door nog meer ponten of door bruggen en tunnels, in de buurt van de Maastunnel kan men meer tunnels bouwen. De fabriek kan zijn reparatiewerkplaats uitbreiden: meer monteurs die naar uitgevallen machines toesnellen. Op vliegvelden kan men extra landingsbanen aanleggen, de loodsdienst kan meer loodsen aantrekken, en de capaciteit van spoorwegemplacementen kan worden vergroot.

De vraag is echter of het verstandig is aan die extra voorzieningen te beginnen. Als bijv. de extra bediende van de kapper, die voorkwam dat U moest wachten, het grootste deel van de tijd niets te doen heeft, is het voor de kapper niet lonend hem in dienst te nemen. Er is dus sprake van een Scylla en een Charybdis. Enerzijds de Scylla van de onvoldoende dienstverlening met als gevolg wachtende mensen of machines of schepen. Anderzijds de Charybdis van het teveel van het goede met als gevolg leegloop bij het dienstverlenend personeel en de bijbehorende installaties.

Het gaat er om op een verstandige manier tussen deze gevaren door te zeilen. De wachttijdtheorie is een onderdeel van de wiskunde dat tot taak heeft in dit opzicht behulpzaam te zijn. De pionier op dit terrein was de Deen A. K. Erlang, die in het begin van deze eeuw zijn onderzoekingen begon als employé van de Kopenhaagse telefoonmaatschappij. Het is niet te verwonderen dat de wachttijdtheorie een betrekkelijk jong vak is, evenmin dat de toepassing op talrijke gebieden een zaak is van nog veel recenter datum. Hildebrand vermeldde dat Stastok Sr. zijn fabriek sloot omdat hij de prijs van zijn product te laag vond; wat vanuit zijn standpunt bezien rationeel was, omdat de installaties en dus de vaste kosten in die tijd niet veel om het lijf hadden. Maar dat is veranderd en tegenwoordig hebben we met kostbare installaties te doen, die door de technische ontwikkeling alleen maar de neiging hebben nog omvangrijker en gecompliceerder te worden. Die te laten stilstaan, geheel of gedeeltelijk, is een kostbare zaak. De betekenis van de wachttijd-*theorie* wordt verder ook verhoogd door de omstandigheid, dat het onmogelijk kan zijn te experimenteren. Men kan er nu eenmaal niet aan beginnen een nieuw vliegveld te bouwen met het uitsluitende doel na te gaan welke invloed dat zal hebben op de tijd, die de vliegtuigen cirkelend moeten doorbrengen voor zij toestemming tot landen krijgen. Voorzover expe-

rimenteren mogelijk is gebeurt dat op papier, d.w.z. door simulatie (een onderwerp dat in het volgende hoofdstuk aan de orde zal komen), maar voordat dit mogelijk is dienen we eerst het nodige te weten over de factoren die een rol spelen en over de wijze waarop zij dit doen. En tenslotte zijn de lonen sinds Stastok aanmerkelijk gestegen, is de werktijd verkort en is de populariteit van ontslag wegens slapte belangrijk minder geworden. Ook leegloop van arbeid is dus meer en meer een probleem geworden.

De variëteit van de toepassingsgebieden maakt het gewenst met een uniforme terminologie te werken. Wij zullen spreken van een *klant* en een *bedieningsinstallatie*. De klant kunt U zijn, maar ook een vliegtuig of een schip. De bedieningsinstallatie kan worden gevormd door een loket, een landingsbaan, een kruispunt, een tunnel, een reparatiewerkplaats. De weg die de klant aflegt kan dan schematisch als volgt worden voorgesteld:

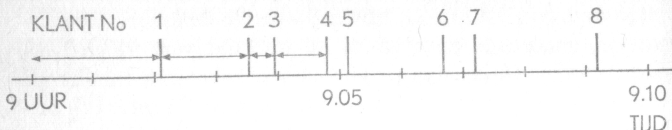
aankomst → (eventueel) wachten → bediening → vertrek

Meestal komen de klanten niet keurig om de 10 minuten (zegge) aan, maar op onzekere tijdstippen. Vaak hebben bovendien niet alle klanten dezelfde bedieningstijd. Het ligt daarom voor de hand de waarschijnlijkheidsrekening te hulp te roepen en de wachttijdtheorie is dan ook, zoals we zullen zien in hetgeen volgt, een onderdeel van de kansrekening met een speciale interpretatie van de begrippen: aankomen, wachten, bedienen.

2. AANKOMST

Wij kunnen het aankomen van de successieve klanten op twee manieren bezien. Wij kunnen ons enerzijds afvragen: hoeveel klanten komen er aan bijv. tussen 9 uur en kwart over 9? Hier gaat het dus om het aantal klanten. Daarvoor komen alleen niet-negatieve, gehele getallen in aanmerking: 0, 1, 2, 3, . . . en de bijbehorende kansverdelingen zijn dan ook van het discrete type. Anderzijds kunnen wij ook de vraag stellen: hoeveel tijd is er verlopen tussen de aankomst van twee successieve klanten? Het antwoord daarop meten we in minuten, seconden, of eventueel tiende seconden. Dat hangt af van de gewenste nauwkeurigheid. In principe hebben we hier te maken met een continue verdeling.

In de volgende grafiek worden deze twee beschouwingswijzen geïllustreerd. Daar is nl. op de tijdas het tijdstip van binnenkomst van een aantal klanten afgezet.

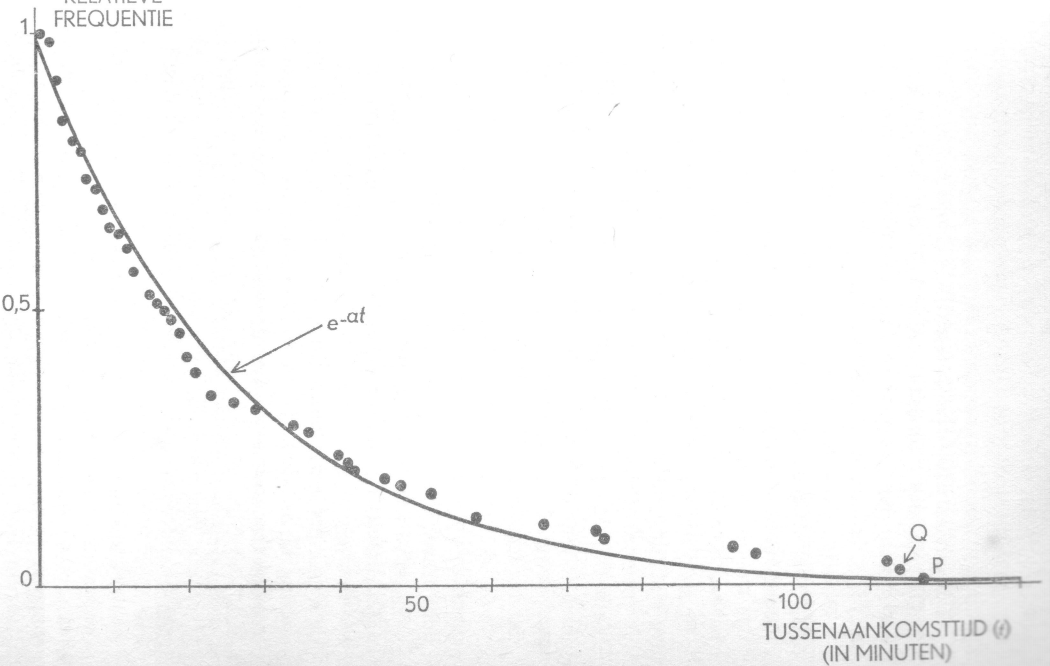


Enerzijds kunnen we aantallen klanten gaan tellen per interval van bijv. 5 minuten. We vinden dan 4 klanten tussen 9 uur en 9.05 en eveneens 4 klanten tussen 9.05 en 9.10. (Dat beide aantallen gelijk zijn is louter toeval. Tussen 9.10 en 9.15 kwamen 7 klanten binnen.) Anderzijds kunnen we de tijd opnemen die verloopt vòòr de eerste klant binnenkomt, en vervolgens de tijd die telkens verloopt tussen het aankomen van twee successieve klanten. We vinden dan, gemeten in seconden:

126 (tot de aankomst van de eerste klant)
85 (vanaf de aankomst van de eerste klant tot de aankomst van de tweede)
26 (vanaf de aankomst van de tweede klant tot de aankomst van de derde)
51 (vanaf de aankomst van de derde klant tot de aankomst van de vierde).

Enzovoorts. Men noemt deze tijdsintervallen tussen twee successieve aankomsten *tussenaankomsttijden.* In de figuur zijn ze met pijlen boven de tijdas aangegeven.

Om het probleem hanteerbaar te maken zullen we iets naders moeten zeggen over de verdeling van de tussenaankomsttijden. Als voorbeeld nemen wij de figuur op blz. 231, die betrekking heeft op meldingen van vliegtuigen bij het radiobaken van Schiphol tijdens de spitsuren van de mistperiode januari-februari 1950. (De gegevens zijn afkomstig van de Rijksluchtvaartdienst.) Horizontaal is de tussenaankomsttijd (t) afgezet; verticaal de relatieve frequentie van het aantal gevallen, waarin de waargenomen tussenaankomsttijd correspondeert met de op de horizontale as aangegeven tijd of nóg groter is. Neem bijv. het meest rechtse punt (P). Dit correspondeert met de grootste waargenomen tussenaankomsttijd (bijna 2 uur verliep toen tussen de aankomsten van twee opeenvolgende vliegtuigen) en er is maar één zo'n geval, dus de verticale coördinaat is de


GECUMULEERDE
RELATIEVE
FREQUENTIE
1
0,5
0
50
100
TUSSENAANKOMSTTIJD (t)
(IN MINUTEN)
$e^{-\alpha t}$
Q
P

relatieve frequentie daarvan $1/N$, waarbij N het totaal aantal waarnemingen voorstelt. Neem vervolgens het punt onmiddellijk linksboven P, dus Q. Dit correspondeert met de een-nagrootste tussenaankomsttijd. Het aantal gevallen, waarin de tussenaankomsttijd daaraan gelijk is of nóg groter, is kennelijk 2. De relatieve frequentie is dus $2/N$, dus is $2/N$ de verticale coördinaat van het punt Q. En zo worden alle punten successievelijk gevonden, gaande naar links en omhoog.

Door de punten is een vloeiende kromme getekend, die er vrij redelijk bij aansluit. De wiskundige gedaante van die kromme is heel eenvoudig, nl. e^{-at}, waarbij $e = 2{,}71828...$, het getal dat we ook al bij de normale verdeling (blz. 175) ontmoet hebben. De a is een coëfficiënt die de vorm van de verdeling in concreto bepaalt. Zijn reciproke, dus $1/a$, is de mathematische verwachting van de tussenaankomsttijd, kortweg: de gemiddelde tussenaankomsttijd. Die zelfde $1/a$ is óók de standaarddeviatie van de tussenaankomsttijden. Dus $1/a$ is zowel de μ als de σ van de verdeling. Er is nog een derde interpretatie: de a zelf, dus niet zijn reciproke, is het aantal klanten dat *gemiddeld* per tijdseenheid aankomt. Om een voorbeeld te nemen, laten we de tijd in minuten meten en laat $a = 3$ zijn. Dan komen er gemiddeld 3 klanten per minuut aan, de gemiddelde tussenaankomsttijd is $^1/_3$ minuut, dus 20 seconden, en de standaarddeviatie van de tussenaankomsttijden is ook 20 seconden.

Dit betreft de zgn. *exponentiële verdeling*, die in de wachttijdtheorie op grote schaal wordt gebruikt om het aspect van de aankomst te hanteren.[1] Hij geeft vaak een heel redelijke aansluiting bij het feitelijk geconstateerde aankomstpatroon. Hij kan ook theoretisch worden afgeleid met behulp van een aantal

1. Het zal de lezer wellicht zijn opgevallen, dat we ons bij de zojuist besproken figuur hebben bezig gehouden met waarden die *tenminste* worden aangenomen, terwijl daarentegen de verdelingsfuncties van Hoofdstuk 6 gebaseerd werden op waarden die *ten hoogste* werden aangenomen. Gemakkelijk ziet men dan in, dat de verdelingsfunctie van de exponentiële verdeling gelijk is aan $1 - e^{-at}$, waarna het voor wie de kunst van het differentiëren machtig is onmiddellijk duidelijk is dat ae^{-at} de dichtheid is. Verwachting en variantie ($1/a$ resp. $1/a^2$) vindt men dan door integreren. De resultaten van de wachttijdtheorie kunnen niet afgeleid worden met de wiskundige middelen waartoe wij ons in dit boek bepalen; zij worden dus zonder bewijs gepresenteerd.

elementaire veronderstellingen, die er op neerkomen dat (vanuit het standpunt van de bedieningsinstallatie) de klanten a.h.w. 'systeemloos' aankomen; maar het zou ons te ver voeren hierop in te gaan. In hetgeen volgt zullen wij het op de exponentiële verdeling van de tussenaankomsttijden houden.

3. BEDIENING

Wanneer een klant is aangekomen zal hij in het algemeen moeten wachten. Dat is dus de volgende fase, maar aangezien het juist gaat om het probleem van het wachten zoals dat door de overige factoren wordt bepaald, zullen wij deze fase in eerste aanleg overslaan. De klant wordt dan geconfronteerd met de volgende vragen:

1. Is er één bedieningsinstallatie of zijn er meer? Er kunnen bijv. verscheidene monteurs in een reparatiewerkplaats zijn, dat komt neer op verscheidene 'bedieningsinstallaties'. We zullen het in eerste instantie eenvoudig houden en ons tot een enkele installatie bepalen.

2. Bestaat er een voorrangsrecht voor bepaalde groepen van klanten of worden zij allen in volgorde van aankomst bediend? Onze oud-vaderlandse molenaars wisten onmiddellijk een antwoord op die vraag, getuige het spreekwoord 'wie het eerst komt, het eerst maalt', en wij zullen ons in eerste aanleg daaraan houden. De voordelen die een voorrangsregeling kan hebben zullen later in dit hoofdstuk aan de orde komen.

En dan, na het wachten, wordt de klant bediend. De *bedieningstijd* is meestal niet dezelfde voor alle klanten en ook hier ligt de introductie van een waarschijnlijkheidsverdeling dus voor de hand. Die verdeling kan uiteraard verschillende vormen hebben. Het blijkt dan dat met name twee karakteristieken van belang zijn, nl. de mathematische verwachting (kortweg de gemiddelde bedieningstijd) en de standaarddeviatie. In het bijzonder de *verhouding* van de standaarddeviatie tot het gemiddelde zal blijken van belang te zijn. Dit quotiënt wordt de 'variatiecoëfficiënt' van de bedieningstijden genoemd. Dus:

$$v = \frac{\text{standaarddeviatie van de bedieningstijden}}{\text{gemiddelde bedieningstijd}}.$$

Zijn de bedieningstijden alle gelijk, dan is hun standaarddeviatie nul en dus ook hun variatiecoëfficiënt v. Zijn de bedieningstijden (net als de tussenaankomsttijden) exponentieel verdeeld, dan hebben we ook een heel speciaal geval. We constateerden immers, dat verwachting en standaarddeviatie van een exponentiële verdeling aan elkaar gelijk zijn, dus is $v = 1$ in dat geval. Merk op, dat we door ons geheel tot verwachting en standaarddeviatie te bepalen veel minder specifiek zijn in onze veronderstellingen t.a.v. de bedieningstijden dan t.a.v. de tussenaankomsttijden. Voor de laatste specificeren wij een volledige verdeling, nl. de exponentiële, en de enige vrijheid die overblijft is de coëfficiënt a, die van geval tot geval kan verschillen. Bij de bedieningstijden gaat het alleen maar om verwachting en standaarddeviatie; het doet er niet toe hoe de verdeling er verder uitziet, althans niet voor het probleem van het gemiddelde aantal wachtende klanten, waarvoor we ons in eerste instantie interesseren. En dat is natuurlijk een voordeel: hoe minder veronderstellingen, hoe beter wat betreft de toepasbaarheid op verschillende concrete situaties.

Het wordt nu tijd, dat we het bedieningsaspect van het probleem en het aankomstaspect aan elkaar koppelen. Want de volgende vraag is van groot belang: kan de bedieningsinstallatie zijn taak aan? Komen er gemiddeld per uur meer klanten binnen dan de installatie kan verwerken, dan wordt de rij wachtenden steeds langer. Er treedt een situatie op, waarin krasse maatregelen nodig zijn, zoals het inrichten van een nieuwe installatie. Anders raken de wachtfaciliteiten verstopt of lopen de klanten weg. Deze catastrofale situatie moet kennelijk uitgesloten worden. Wiskundig bezien komt dit neer op een vergelijking van de gemiddelde tussenaankomsttijd ($1/a$) met de gemiddelde bedieningstijd. Is de laatste groter, dan loopt het mis. Dan is er meer tijd nodig om de klanten te helpen dan die klanten nodig hebben om aan te komen, althans gemiddeld genomen, en de rij wachtenden gaat voortdurend langer worden. Is de gemiddelde bedieningstijd kleiner, dan is die zorg er niet. Beschouw nu de volgende verhouding:

$$b = \frac{\text{gemiddelde bedieningstijd}}{\text{gemiddelde tussenaankomsttijd}} = \frac{\text{gemiddelde bedieningstijd}}{1/a}$$
$$= a \times \text{gemiddelde bedieningstijd.}$$

Deze verhouding moet dus kleiner dan 1 zijn opdat het systeem niet verstopt raakt. Hij staat bekend als de *bezettingsgraad* van de bedieningsinstallatie. Is bijv. $b = 0{,}7$, dan betekent dit dat de gemiddelde tijd om een klant te bedienen 70% bedraagt van de tijd die gemiddeld verstrijkt tussen de aankomst van twee opeenvolgende klanten; maar het betekent óók, naar valt aan te tonen, dat de installatie gemiddeld genomen 70% van de tijd bezet is, dus met een klant bezig is, en in de overige 30% zit te wachten op een nieuwe klant. Vandaar de term 'bezettingsgraad'.

4. DE GEMIDDELDE RIJLENGTE

We hebben nu onze bouwstenen gereed om het eigenlijke probleem aan te pakken: het wachten. Maar we moeten even in de stijl van het probleem blijven door aan een voorafgaande opmerking voorrang te geven. In het begin van dit hoofdstuk hadden we het over een loket dat om half negen open ging, en we constateerden dat de eerste klant onmiddellijk wordt geholpen, dus helemaal niet hoeft te wachten. (We nemen aan dat hij er niet vóór half negen is gaan staan.) De tweede klant behoeft niet te wachten als hij na het vertrek van de eerste arriveert, maar het kan ook zijn dat hij wél moet wachten, nl. wanneer hij komt als de eerste klant er nog staat. Kennelijk is de situatie voor de tweede klant verschillend van die van de eerste. Komt de derde klant binnen terwijl de tweede achter de eerste staat te wachten, dan staan er twee in de rij te wachten. Ook dat is een nieuwe situatie, want twee wachtende klanten is een ondenkbare situatie voordat de derde klant binnenkomt. En zo gaat het door. Het punt is nu, dat als de bezettingsgraad kleiner dan 1 is, dit soort verschillen geleidelijk aan gaat verdwijnen naarmate de tijd voortschrijdt, dus naarmate er meer klanten aankomen. Dat betekent *niet*, dat èlke klant dan even lang moet wachten. Integendeel, sommigen zullen lang moeten wachten, sommigen korter. De tijd die de klanten wachtend moeten doorbrengen is aan een verdeling onderhevig, zoals men onmiddellijk zal aanvoelen; de tijdstippen van aankomst zijn immers onregelmatig en de bedieningstijden zijn verschillend, dus het kan niet anders dan dat de wachttijden ook verschillend zijn. Maar het betekent wèl, dat de *verdeling* hetzelfde gaat worden; d.w.z. de 100ste klant mag misschien tweemaal zo lang wachten als de 200ste, maar dat berust niet op syste-

matische gronden, alleen maar op het feit dat de verdeling van de wachttijd nu eenmaal door een zekere spreiding is gekenmerkt. Bij de eerste klant is dat bepaald nog niet het geval; wij zagen immers dat zijn wachttijd nul is, en wel met zekerheid. Voor de tweede klant geldt dat hij met positieve waarschijnlijkheid moet wachten, zijn wachttijdverdeling is dus anders. Maar geleidelijk aan raakt, zoals dat heet, het systeem in een 'stationnaire toestand' en de wachttijdverdeling is dan een en dezelfde voor alle successieve klanten. Hoe lang het duurt voor die toestand bereikt is hangt in sterke mate af van de bezettingsgraad. Is die ver beneden 1, dan gaat het vlug; ligt die dicht bij 1, dan kan er een hele tijd over heengaan voor we werkelijk in de buurt van de stationnaire toestand zitten. In hetgeen volgt zullen wij ons eenvoudigheidshalve bepalen tot de situatie, waarin die toestand bereikt is.

Nu dan het wachten. Men ziet gemakkelijk in, dat dit probleem van twee kanten benaderd kan worden, precies als bij de aankomst het geval was. Daar konden we hetzij het aantal klanten tellen dat per tijdseenheid aankomt, hetzij de tijd meten die verstrijkt tussen de aankomst van twee opeenvolgende klanten. Hier, bij het wachten, kunnen we hetzij het aantal klanten tellen dat staat te wachten achter de klant die bediend wordt, hetzij de tijd meten die een klant wachtende doorbrengt. Wij zullen de eerste aanpak volgen, dus die van de rijlengte. Dat betreft dus 0, 1, 2, . . . klanten die staan te wachten; de verdeling van de rijlengte is kennelijk discreet. Wij zullen ons bezig houden met de verwachting van die verdeling, dus met de *gemiddelde rijlengte*, r. Onder de gemaakte veronderstellingen (exponentieel verdeelde tussenaankomsttijden en stationnaire toestand) valt aan te tonen, dat die gemiddelde rijlengte bepaald wordt door slechts twee factoren: de bezettingsgraad b en de variatiecoëfficiënt van de bedieningstijden, v. De formule luidt als volgt:

$$r = \frac{b(1 - {}^1/_2 b) + {}^1/_2 b^2 v^2}{1 - b}.$$

We zien dat, wanneer v toeneemt bij constante b, ook r toeneemt. Een grote spreiding in de bedieningstijden is dus ongunstig voor de gemiddelde rijlengte! We konden dit verwachten. Want naarmate deze spreiding groter is komt er vaker een klant binnen waarvan de bediening bijzonder veel tijd in beslag neemt. Het aantal wachtenden kan in die tijd hoog oplo-

pen. Weliswaar komt er dan ook vaak een klant binnen die heel weinig tijd nodig heeft, hetgeen de rijlengte reduceert. Maar zo'n reductie bedraagt iedere keer maar één klant en de vergroting van de rijlengte door een tijdrovende klant kan heel wat omvangrijker zijn.

Het verband tussen b en r is iets moeilijker uit de formule te halen. Daarom is de volgende tabel instructief. Hij bevat de gemiddelde rijlengte voor verschillende waarden van de bezettingsgraad b en de variatiecoëfficiënt van de bedieningstijden v:

	$b=0,5$	$b=0,7$	$b=0,85$	$b=0,95$	$b=0,98$	$b=0,99$
$v=0$	0,75	1,51	3,26	10	25	50
$v=1/2$	0,81	1,72	3,86	12	31	62
$v=1$	1,00	2,33	5,67	19	49	99
$v=2$	1,75	4,78	13	46	121	246
$v=4$	4,75	15	42	154	409	834

Wanneer we verticaal door de tabel gaan zien we een duidelijke stijging van de gemiddelde rijlengte naarmate v stijgt, geheel in overeenstemming met de conclusie die we reeds uit de formule getrokken hadden. Bij een bezettingsgraad van 95% bijv. ($b = 0,95$) zien we dat er gemiddeld 10 klanten staan te wachten wanneer alle klanten dezelfde bedieningstijd hebben ($v = 0$); en dat loopt op tot een gemiddelde van 19 wachtende klanten, dus bijna het dubbele, als de standaarddeviatie van de bedieningstijden even groot is als hun gemiddelde ($v = 1$), zoals bij exponentieel verdeelde bedieningstijden het geval is. Gaan we vervolgens van links naar rechts door de tabel, dan zien we boven $b = 0,85$ en met name wanneer b nog dichter in de buurt van de 1 komt, de gemiddelde rijlengte scherp stijgen. De noemer van de formule ($1 — b$) komt in dat geval namelijk zeer dicht bij nul te liggen, met als gevolg een explosieve situatie. Rechtsonder in de tabel ziet U 834 klanten van ongeduld trappelen (en dat is nog maar een gemiddelde!). Natuurlijk is dit geen realistische situatie; tegen die tijd is het systeem al lang ter ziele en moet het door iets verstandigers vervangen worden. Maar wat de tabel wel leert is dat kleine verschuivingen grote gevolgen kunnen hebben. Als de landingsfaciliteiten op Schiphol bijv. een bezettingsgraad van 95% zouden hebben en er komen nog een paar vliegtuigen per dag bij, dan stijgt de bezettingsgraad allicht tot bijv. 99% en het aantal rondcirkelende vliegtuigen stijgt dan tot het vijfvoudige of meer!

5. VOORRANG

Wij veranderen nu een van onze basisveronderstellingen: we gaan er niet langer van uit, dat alle klanten zonder onderscheid naar volgorde van aankomst worden bediend, maar we houden nu rekening met prioriteiten. Dat kennen we in het dagelijks leven op grote schaal: we laten dames voorgaan omdat we hoffelijk willen zijn en in de meer zakelijke sfeer wordt bij P.T.T.-loketten voorrang verleend aan wie een telegram komt aanbieden. De redenen van prioriteitsverlening kunnen dus verschillend zijn. Bij de reparatie van defecte machines kan aan een bepaald exemplaar voorrang worden verleend omdat het in een spoedkarwei is betrokken of omdat het een delicaat exemplaar is dat bij verder uitstel dreigt door te roesten, enz. Hier zullen wij ons bezig houden met prioriteitsverlening die erop gericht is de gemiddelde wachttijd te bekorten. Dit onderwerp kan op aardige wijze geïntroduceerd worden met het volgende historische geval. De jeugd had een brandje gesticht en de bromfiets van de heer B. was daardoor nogal ernstig beschadigd. De reparatie zou nogal tijdrovend zijn. Toen de heer B. na enige tijd bij zijn bromfietshersteller informeerde hoe de toestand was, kreeg hij ten antwoord: 'Meneer, ik doe altijd eerst de kleine klussies, want ik heb liever één ontevreden klant dan twintig.' Nu zult U wellicht zeggen: allicht, 20 is groter dan 1, maar de heer B. is dan ook een dure klant – hij levert heel wat meer geld op dan elk van die 20. En het punt is dan dit: zelfs al zou de bromfietshersteller daarmee rekening houden door de belangen van de heer B. te 'wegen' met het bedrag dat hij dadelijk van hem zal incasseren (gebaseerd op een vast uurloon), zelfs dan doet hij er verstandig aan eerst voor de 20 andere klanten te gaan werken. Men kan nl. de gemiddelde wachttijd van de groep van *alle* klanten *in totaal* reduceren door prioriteit te geven aan hen die een korte bedieningstijd hebben.

Hier gaan wij nu meer in detail op in. We stellen ons voor, dat de klanten in twee groepen worden opgedeeld. De eerste groep omvat de 'snelle' klanten, de tweede de tijdrovende. De gemiddelde bedieningstijd van de eerste groep is dan kleiner dan die van de tweede; we schrijven q voor hun verhouding:

$$q = \frac{\text{gemiddelde bedieningstijd van de snelle groep}}{\text{gemiddelde bedieningstijd van de tijdrovende groep}}$$

Kennelijk is q een getal tussen 0 en 1. Bijv., laten de snelle klanten een gemiddelde bedieningstijd van 6 hebben (in minuten, uren of wat ook) en de tijdrovende van 24, dan is $q = 6 / 24 = 0{,}25$. Dit is het enige dat wij voor ons resultaat nodig zullen hebben wat betreft de bedieningstijden van de twee groepen. In het algemeen zullen beide groepen door een zekere spreiding in de bedieningstijd gekenmerkt zijn en het is dus best denkbaar, dat de snelle groep een aantal klanten bevat, die bij bediening blijken tijdrovender te zijn dan sommige klanten van de tweede groep.

Vervolgens dienen we ons bezig te houden met de samenstelling van de klanten naar de twee groepen. Hoeveel procent behoort tot de ene groep, hoeveel tot de andere? Ook dit beschouwen we als een gegeven voor ons probleem en we schrijven l voor de fractie van de klanten in de eerste groep. Dus:

$100l$ = aantal klanten van de snelle groep, berekend als percentage van het totaal.

Als de eerste groep dus 20% van de klanten omvat, hebben we $l = 0{,}2$ en een fractie $1 - l = 0{,}8$ van klanten van de tijdrovende groep.

Hiermee hebben we de begrippen omschreven die we nodig hebben i.v.m. de gewijzigde probleemstelling: het onderscheid naar de twee groepen. Hoe werkt nu de voorrangsprocedure? We gaan van het volgende uit. Zodra een klant vertrekt wordt zijn plaats bij de bedieningsinstallatie ingenomen door de eerst binnengekomen klant van de eerste groep. Pas als er geen klanten van de eerste groep staan te wachten komt de eerst aangekomen klant van de tweede groep aan de beurt. (Binnen de twee groepen afzonderlijk gaan we dus te werk volgens de volgorde van aankomst.) Is een klant van de tijdrovende groep eenmaal onder behandeling, dan wordt die behandeling tot het einde voortgezet en dus niet onderbroken door de aankomst van een klant van de eerste groep. Dat laatste kan ook, maar met dat geval zullen we ons niet bezig houden.

Het is niet moeilijk in te zien, dat deze voorrangsprocedure moet 'werken' in de zin dat de totale wachttijd voor alle klanten tezamen, en dus ook de gemiddelde wachttijd per klant voor de groep van alle klanten, wordt gereduceerd. Neem het eenvoudige geval van twee klanten, A en B, die in die volgorde

binnenkomen. De bedieningsinstallatie is bezet, beiden moeten wachten. Laat *A* tot de tweede groep behoren en *B* tot de eerste en laten hun bedieningstijden 20 resp. 8 minuten zijn. Eerst de procedure van strikte behandeling naar volgorde van aankomst. Dan wordt *A* eerst geholpen en vanaf het moment waarop dat gaat gebeuren heeft *B* 20 minuten te wachten. Na afloop daarvan vertrekt *A*, en *B* wordt geholpen; nog 8 minuten later is hij weg, in totaal dus 28 minuten. De totale wachttijd voor beiden (gemeten vanaf het moment waarop *A* wordt bediend) is 20 minuten, geheel voor rekening van B. Vervolgens de prioriteitsprocedure. Dan wordt *B* eerst geholpen (hij behoort tot de eerste groep, *A* tot de tweede) en hij vertrekt na 8 minuten. Die 8 minuten moet *A* wachten en dan komt hij aan de beurt (tenzij in de tussentijd een andere klant van de eerste groep aankomt) en na nog 20 minuten is ook hij weg; in totaal weer 28 minuten. Maar nu wordt er in totaal slechts 8 minuten gewacht, ditmaal voor rekening van *A*. De prioriteitsprocedure leidt dus tot een 'winst' van 20 — 8 = 12 minuten in totaal. Uiteraard is er geen winst voor *A*, integendeel; de groep die prioriteit verleent moet langer wachten en de andere groep korter, dat compenseert elkaar gedeeltelijk, maar per saldo blijft er een winst van 12 minuten over. Merk op dat die verwisseling geen gevolgen heeft voor de klanten die na *B* binnenkomen. *A* en *B* kosten tezamen 28 minuten, dat geldt onafhankelijk van de vraag wie het eerst bediend wordt.

Nu het algemene geval. We zullen ons bezig houden met de procentuele reductie van de gemiddelde wachttijd (voor de groep van alle klanten in totaal), die het gevolg is van de prioriteitsverlening. Die blijkt bepaald te zijn door drie factoren: de verhouding van de gemiddelde bedieningstijden, q; de fractie van de klanten behorend tot de groep die prioriteit ontvangt, l, en de bezettingsgraad b. Voor b kunnen we terugverwijzen naar de voorgaande bladzijden; de mate waarin de installatie bezet is verandert niet door de invoering van een prioriteitsprocedure, want die doet niets anders dan klanten van plaats verwisselen. Het resultaat is als volgt:

$$\text{Procentuele reductie van de gemiddelde wachttijd} = \frac{100l(1-l)(1-q)b}{1-l+lq(1-b)}.$$

Een voorbeeld is instructief. Neem $q = 0{,}3$, zodat de eerste

groep een gemiddelde bedieningstijd heeft die 30% bedraagt van die van de tijdrovende groep. Voor dat geval kunnen we de afhankelijkheid van de procentuele reductie van de twee andere bepalende factoren, *l* en *b*, uit de volgende tabel aflezen:

	$b = 0{,}5$	$b = 0{,}7$	$b = 0{,}85$	$b = 0{,}95$	$b = 0{,}99$
$l = 0{,}1$	3	5	6	7	7
$l = 0{,}2$	7	10	12	13	14
$l = 0{,}3$	10	14	18	20	21
$l = 0{,}4$	13	18	23	26	28
$l = 0{,}5$	15	22	28	33	35
$l = 0{,}6$	17	26	33	39	41
$l = 0{,}7$	18	28	38	45	48
$l = 0{,}8$	$17^1/_2$	29	40	50	55
$l = 0{,}9$	13	24	38	53	61

Wanneer aan slechts weinig klanten voorrang wordt verleend is de reductie natuurlijk bescheiden. In de eerste rij van de tabel ($l = 0{,}1$, dus aan 10% van de klanten wordt prioriteit verleend) varieert de reductie tussen 3 en 7 procent. Loopt *l* op, dan kan de mogelijke reductie ook belangrijk toenemen, maar tegen de tijd dat bijna iedereen voorrang krijgt ($l = 0{,}9$) kan het rendement van de maatregel weer wat teruglopen. Dit laatste geldt i.h.b. wanneer de bezettingsgraad niet te hoog is. Wat de invloed van deze factor betreft, we zien dat er vooral bij hoge bezettingsgraad mogelijkheden van aanzienlijke reducties zijn: de gemiddelde wachttijd kan bij bezettingsgraden van 95% en hoger met ongeveer een derde worden bekort wanneer de eerste groep even groot is als de tweede ($l = 0{,}5$) en met bijna de helft wanneer de eerste groep 70% van de klanten omvat. Dit is vooral daarom zo belangrijk, omdat die wachttijden bij hoge bezettingsgraden zonder prioriteitsprocedure extreem hoog kunnen zijn. We hebben daarvan het een en ander gezien toen we de gemiddelde rijlengte bestudeerden en het laat zich raden hoe dit doorwerkt voor de gemiddelde wachttijd.

Er zijn vele situaties denkbaar waarbij een reductie van de gemiddelde wachttijd leidt tot een evenredige kostenbesparing. Besparingen van meer dan 25 procent, zoals erin de tabel ettelijke voorkomen, zijn dan bepaald wel de moeite waard. 'Wie het eerst komt, het eerst maalt' is dus niet in alle situaties de hoogste wijsheid.

We zullen nu een concreet geval nader bezien. In een fabriek staat een aantal automatische machines. Wanneer alles goed gaat behoeven ze niet door mensenhand bediend te worden; pas wanneer er iets misloopt moet er iemand naar kijken. Nu is het in de regel niet te voorspellen wanneer een machine zal uitvallen, het kan zich ieder ogenblik voordoen. Er is dus alle reden van de waarschijnlijkheidsrekening gebruik te maken. We zullen aannemen dat de looptijd van een machine, d.i. de tijd die er verloopt tussen het tijdstip van starten en van (opnieuw) uitvallen van een machine, exponentieel verdeeld is met gemiddelde $1/a$. Meten we de tijd in uren en is $a = 0{,}2$, dan is de gemiddelde looptijd van zo'n machine dus 5 uur. Merk op dat dit vergelijkbaar is met een tussenaankomsttijd, waarvoor we van meet af aan met een exponentiële verdeling hebben gewerkt. Iedere keer nl., als de looptijd over is, komt de machine 'terug' om gerepareerd te worden.

Vervolgens de 'bediening', in dit geval de reparatie van de machine door een monteur. Wij gaan ervan uit, dat er de ene keer meer aan de machine kan mankeren dan de andere keer, dus dat de reparatietijden verschillend kunnen zijn. Preciezer, we veronderstellen dat de reparatietijd ook exponentieel verdeeld is en wel met een gemiddelde dat 10 maal zo klein is als dat van de looptijd, dus $1/10a$. Is de gemiddelde looptijd 5 uur, dan is de gemiddelde reparatietijd dus een half uur. En tenslotte, wat de aantallen betreft: we nemen het geval van zes machines en één monteur.

Er zijn nu zeven mogelijkheden, afhankelijk van het aantal defecte machines (0, 1, 2, . . ., 6). In de eerste plaats kunnen alle machines naar tevredenheid werken, zodat de monteur niets te doen heeft en ook geen enkele machine op reparatie staat te wachten. Onder de hierboven opgesomde veronderstellingen is de kans dat dit gebeurt, zoals na pittig rekenen[1] blijkt, gelijk aan 0,48. Dat betekent dus, dat de monteur in bijna de helft van de tijd geen emplooi heeft aan de machines (behoudens poetsen e.d.). In de tweede plaats kan 1 machine uitvallen. De monteur snelt daar op af en er is nog steeds geen machine die op reparatie staat te wachten. Deze mogelijkheid heeft een kans van

1. En onder de veronderstelling dat de stationaire toestand is bereikt.

0,29. In de derde plaats kunnen 2 machines uitvallen. De monteur kan er maar één helpen, 1 machine staat op reparatie te wachten. Deze mogelijkheid heeft een kans van 0,15. Zo gaan we door tot aan de mogelijkheid van het gelijktijdig defect zijn van alle machines (de kans daarop is maar heel klein); iedere keer is het aantal wachtende machines één minder dan het aantal dat defect is. Een volledig overzicht vindt men hieronder:

Aantal machines dat *niet* werkt	Aantal machines dat moet wachten	Waarschijnlijkheid, dat deze situatie optreedt
0	0	0,48
1	0	0,29
2	1	0,15
3	2	0,06
4	3	0,02
5	4	0,0035
6	5	0,0003
	Totaal:	1

Zoals wij zagen is 0,48 de kans dat de monteur zit te wachten op het defect raken van een machine. Daar staat tegenover, dat van tijd tot tijd machines wachten op een monteur. Gemakkelijk verifieert men, dat de mathematische verwachting van het aantal wachtende machines gelijk is aan:

$$0{,}48 \times 0 + 0{,}29 \times 0 + 0{,}15 \times 1 + \ldots + 0{,}0003 \times 5 = 0{,}33.$$

Op dezelfde wijze kunnen we het geval van 20 machines en 3 monteurs beschouwen. Het gaat om machines van hetzelfde type als in het vorige voorbeeld, en om monteurs van dezelfde graad van bekwaamheid en met hetzelfde werktempo. De verdelingen van de looptijden van de machines en van de bedieningstijden zijn dus dezelfde als voorheen. De situaties die zich nu kunnen voordoen zijn opgesomd in het overzicht op blz. 244. De eerste regel komt overeen met de situatie dat alle machines lopen. Het aantal machines dat *niet* werkt is dus nul, géén machine wordt gerepareerd, géén machine wacht, de drie monteurs hebben geen werk. Deze situatie blijkt zich – onder de boven gegeven voorwaarden – in 14 procent van de tijd voor te doen. Valt er 1 machine uit (tweede regel), dan begint er onmid-

Aantal machines dat *niet* werkt	Aantal machines in reparatie	Aantal machines dat moet wachten	Aantal monteurs dat niets te doen heeft	Waarschijnlijkheid, dat deze situatie optreedt
0	0	0	3	0,14
1	1	0	2	0,27
2	2	0	1	0,26
3	3	0	0	0,16
4	3	1	0	0,09
5	3	2	0	0,05
6	3	3	0	0,02
7	3	4	0	0,011
8	3	5	0	0,005
9	3	6	0	0,002
10	3	7	0	0,0007
11	3	8	0	0,0002
12	3	9	0	0,00007

dellijk een monteur aan te werken, géén machine wacht, de twee andere monteurs hebben niets te doen. Kans: 0,27. Vallen 2 machines uit (kans 0,26), dan is er maar één monteur die niets te doen heeft. Bij 3 defecte machines (kans 0,16) hebben allen werk en nog steeds is er geen wachtende machine. Dat gaat beginnen wanneer 4 machines tegelijkertijd defect zijn en van dan af aan is het aantal wachtende machines steeds 3 minder dan het aantal dat defect is. De tabel eindigt bij 12 defecte machines; een groter aantal kan ook (tot alle 20 toe), maar de kans daarop is zo klein dat die gevallen niet interessant zijn.

Hoe vergelijken we nu die twee gevallen? Enerzijds 1 monteur op 6 machines, anderzijds 3 monteurs op 20 machines. Laten we eerst de leegloop van monteurs bezien. De ene monteur van het eerste geval had een leegloop van 48%. In het tweede geval is de mathematische verwachting van het aantal monteurs dat niets te doen heeft:

$$0{,}14 \times 3 + 0{,}27 \times 2 + 0{,}26 \times 1 = 1{,}21.$$

Aangezien er drie monteurs zijn is dat 1,21 / 3 = 0,40 per monteur. Dat is kennelijk beter dan de 0,48 van het eerste geval. Allicht, zult U misschien zeggen; in het tweede geval is de monteur-machineverhouding 1 op $6^2/_3$, dat is minder dan de 1 op 6 van het eerste geval en dus hebben de monteurs meer te doen.

Het zal wel anders zijn, zo zult U wellicht vervolgen, met het wachten van de machines, want die moeten in het tweede geval met minder monteurs tevreden zijn.

Dat zullen wij nu bezien. Wij vonden dat in het eerste geval de mathematische verwachting van het aantal wachtende machines 0,33 is. Er zijn er zes in dat geval, dus per machine is dat $0{,}33 / 6 = 0{,}055$ of $5^1/_2\%$. In het tweede geval is de verwachting:

$$0{,}09 \times 1 + 0{,}05 \times 2 + 0{,}02 \times 3 + \ldots = 0{,}34.$$

Per machine is dat $0{,}34 / 20 = 0{,}017$ of 1,7%. Dus ruim drie keer zo laag! Waarom is het zo, dat het machineverlies hier geringer is ondanks het relatief geringere aantal monteurs? Dat is omdat bij die grotere aantallen de verhoudingen beter komen te liggen. En dat is een van de voordelen van het grootbedrijf!

LITERATUUR

Een zeer eenvoudig en plezierig geschreven inleiding is het recente artikel van Sittig [1]. Wil men dieper op het onderwerp ingaan, dan kan men terecht bij het boek van Morse [2] en bij Hoofdstuk 11 van dat van Saaty [3]. Het voorrangsmodel werd voor het eerst bestudeerd door Cobham [4]; de op blz. 237 en 241 afgedrukte tabellen zijn uittreksels uit grotere tabellen, gepubliceerd door J. Koerts [5]. De tabellen op blz. 243 en 244 zijn oorspronkelijk afkomstig uit een door C. Palm in het Zweeds geschreven artikel in het tijdschrift *Industritidningen Norden* van 1947. Palm geeft tabellen en grafieken om het optimale aantal monteurs te bepalen. In het Engels kan men erover lezen in Feller's leerboek [6, pp. 416-420].

[1] Sittig, J., 'Enige opmerkingen over wachttijdproblemen'. *Literatuuroverzicht van het Technisch Documentatie- en informatiecentrum voor de Krijgsmacht*, Jaargang 6 (1962), pp. B72-93.

[2] Morse, P. M., *Queues, Inventories and Maintenance.* John Wiley and Sons, Inc., New York. 1958.

[3] Saaty, T. L., *Mathematical Methods of Operations Research.* McGraw-Hill Book Company, Inc., New York. 1959.

[4] Cobham, A., 'Priority Assignment in Waiting Line Problems'. *Operations Research*, Vol. 2 (1954), pp. 70-76 [met een correctie in Vol. 3 (1955), p. 547].

[5] Koerts, J., 'On Mean Waiting Times and Their Reduction by Priority Procedures: An Expository Survey and Some Tables'. *Statistica Neerlandica*, Jaargang 17 (1963), pp. 267-283.

[6] Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. I. Second Edition. John Wiley and Sons, Inc., New York. 1957.

10. SIMULATIE EN BELEIDSSPELEN

1. DE VOETBALPOOL

Iedere week wordt er door ongeveer een half miljoen Nederlanders gewed op de resultaten van voetbaluitslagen: de voetbalpool. Dit spel munt uit door eenvoud. Op een formulier staan 15 wedstrijden vermeld, waarvan de resultaten moeten worden voorspeld. Het is echter niet nodig de precieze score aan te geven. Voldoende is het te voorspellen welke club zal winnen, of te voorspellen dat het een gelijk spel zal worden. Dit doet men door het invullen van een 1 (de thuisclub wint), een 2 (de bezoekende club wint) of een 3 (het wordt een gelijk spel). Zo heeft men na invulling een rijtje van 15 cijfers, alle gelijk aan 1, 2, of 3. Bijvoorbeeld:

1 2 1 3 3 1 2 1 1 2 2 1 1 3 1

Dus inderdaad heel eenvoudig. We laten details betreffende reservewedstrijden, 'treffertips' en extra beloningen wegens veel goed voorspelde gelijke spelen buiten beschouwing.

De volgende zondag worden de 15 wedstrijden gespeeld. De uitslagen hiervan worden nog diezelfde middag door de radio vermeld, zodat iedereen kan controleren of hij juist heeft voorspeld. In welk geval men een flinke prijs kan winnen. In de regel echter zal men zijn hoop op de volgende week moeten richten, omdat er maar 6, of misschien 9, van de 15 wedstrijden goed werden voorspeld. Zelfs de experts in de kranten plegen nauwelijks de helft van de wedstrijden correct te voorspellen (gemiddeld genomen).

Ongelukkigerwijze is er één probleem bij de organisatie. De weersomstandigheden zijn soms zo bizar slecht dat de wedstrijden de volgende zondag *niet* worden gespeeld. Het terrein is door dagenlang onophoudelijk regenen geheel doorweekt, of een lange vorstperiode heeft de velden een soort cement-structuur gegeven. De wedstrijden worden dan afgelast, en men kan niet controleren of men juist heeft voorspeld om de eenvoudige reden dat de wedstrijd niet is gespeeld. Wat dan? U denkt wellicht dat men dan maar moet wachten totdat, tenslotte, alle wedstrijden wèl zijn gespeeld. Doch deze handelwijze heeft prac-

tische bezwaren: de voorspellingspapiertjes moeten lang worden bewaard en kansen op fraude nemen toe. En bovendien theoretische bezwaren: de voorspellingen golden de wedstrijden voor a.s. zondag; als men ze later speelt zijn de omstandigheden en gegevens waarop de voorspelling was gebaseerd in de regel nogal gewijzigd. Deze handelwijze wordt daarom niet gevolgd.

Eenvoudiger is het de pool-formulieren ongeldig te verklaren, en de prijzenpot bij die van de volgende week op te tellen. Dit gebeurt in Nederland, maar niet in Engeland. Daar heeft men het probleem wat krachtiger aangepakt. Als dáár de wedstrijden niet worden gespeeld, dan worden zij *gesimuleerd*. Althans, het resultaat wordt gesimuleerd. Dit gebeurt op een heel interessante manier. Vijf experts (bijv. scheidsrechters, spelers of clubbestuurders) verschijnen tezamen op het televisiescherm, en bespreken daar de kansen voor de teams in een bepaalde, niet gespeelde, wedstrijd: bijv. Arsenal-Blackpool. Zij houden daarbij rekening met een groot aantal relevant veronderstelde omstandigheden: resultaten in het verleden, ziekte van spelers e.d. Naar aanleiding van deze discussie zou de voorzitter bijvoorbeeld kunnen concluderen dat 'alles welbeschouwd, de thuisclub Arsenal een kans had om te winnen van 0,55 en een kans om te verliezen van slechts 0,10'. De kans op een gelijk spel ligt daarmee vast op 0,35, want één van de drie uitslagen zou zeker zijn geregistreerd als de wedstrijd zou zijn gespeeld.

De *kansen* op een 1 (winst voor Arsenal), 2 (winst voor Blackpool) en een 3 (gelijk spel) zijn hiermee bepaald op laten we zeggen een redelijk objectieve wijze, in elk geval volkomen in het openbaar. Nu komt het écht interessante. Er verschijnt een grote jute zak met 100 golfballen, genummerd van 1 tot en met 100. Het ligt nu voor de hand om deze ballen op te splitsen in drie groepen, groep 1 genummerd van 1 t/m 55, groep 2 genummerd van 56 t/m 65, en groep 3 genummerd van 66 t/m 100. Deze groepen bevatten respectievelijk 55, 10 en 35 ballen – dus precies dezelfde verhouding als de vastgestelde kansen op een resultaat 1, 2 of 3! De voorzitter trekt nu een bal uit de zak waarin alle 100 ballen goed dooreengeschud voorkomen. Hij leest het nummer af en laat dit bovendien op het scherm zien. Als het één der getallen 1 t/m 55 is, geldt een '1' als de juiste voorspelling; een '2' wordt het resultaat als het getal 56 t/m 65, dus bijv. 61 is. Tenslotte, als het getrokken nummer groter is dan 65, wordt een 3 als correcte voorspelling beschouwd. (Om

misverstand te voorkomen voegen we hieraan toe, dat het op deze wijze gesimuleerde resultaat alléén voor de poolformulieren van belang is; niet bijvoorbeeld voor de competitiestanden.)

Het lijkt misschien onredelijk dat als nummer 61 wordt getrokken, Blackpool (op de poolformulieren) heeft gewonnen, want ze hadden maar zo'n kleine kans, slechts 1 op 10. Maar dit is een logisch gevolg van de bewering *dat* ze een kans hebben van 1 op 10. Als ze die kans hebben, moeten ze die ook krijgen! Het zou bepaald onjuist zijn te verlangen dat men een 1 invult zonder de andere mogelijke resultaten een kans te geven alleen omdat het 't waarschijnlijkst wordt geacht dat Arsenal zal winnen.

Al met al hebben we hiermee een eenvoudig, maar actueel en realistisch voorbeeld gegeven van simulatie. Wij onderkennen hierbij het *model*, d.w.z. een veronderstelde afbeelding van de werkelijkheid; in dit geval bestaat dat eenvoudig uit de vastgestelde kansen op de resultaten 1, 2 of 3; dus bij Arsenal-Blackpool respectievelijk 0,55 en 0,10 en 0,35. En daarnaast hebben we een *mechanisme* om het model na te bootsen. In dit geval een mechanisme om gebeurtenissen 1, 2 en 3 met kansen 0,55 resp. 0,10 resp. 0,35 te 'genereren'. Dit mechanisme bestond uit een zak met 100 genummerde golfballen. Deze twee karakteristieken, een model en een mechanisme om het model na te bootsen, zijn noodzakelijk voor simulatie. Het testen van scheepsmodellen in waterbakken, waarin kunstmatig golven en stromingen worden opgewekt, is een ander voorbeeld van simulatie. Men kan op deze wijze het scheepsmodel toetsen op stabiliteit, wendbaarheid, e.d. Men doet dat hetzij om de gemaakte berekeningen te controleren, ofwel omdat het te ingewikkeld is die berekeningen uit te voeren.

2. EEN MARGARINEFABRIEK

Reeds in Hoofdstuk 9 over wachttijdproblemen werd gesproken over een machinepark. De machines konden stuk gaan, en dan moest er een monteur bij te pas komen. Er waren twee onaangename situaties: de monteur had niets te doen, of een defecte machine moest op een monteur wachten. Beide omstandigheden kosten geld. De vraag is, hoe deze kosten te minimeren, gegeven zekere kansuitspraken

over het defect raken van de machines en hun reparatieduur.

Nu werd in het hoofdstuk over wachttijden gesteld, dat als men enkele veronderstellingen maakt betreffende het optreden van defecten en de benodigde reparatieduur – bijv. dat beide exponentieel verdeeld zijn –, dit probleem wiskundig exact is op te lossen. De berekeningen werden niet in detail vermeld, en wellicht werd de indruk gewekt dat dit niet eens zo'n heksentoer was. Nu, het probleem was dan ook heel eenvoudig, en de veronderstellingen waren met zorg zo gekozen dat zij voor wiskundige manipulaties vatbaar waren. (Het is een van de kleine wondertjes van de natuur dat deze veronderstellingen in de regel merkwaardig realistisch zijn en redelijk nauwkeurige resultaten geven.) Maar als de problemen wat ingewikkelder worden, wordt de vereiste wiskunde ook ingewikkelder. Nog geheel afgezien van de mogelijkheid dat de wiskundige oplossing niet te geven is omdat de veronderstellingen wiskundig onhandelbaar zijn, of omdat te veel factoren tegelijkertijd een rol spelen. Zelfs als een oplossing wel langs wiskundige weg kan worden afgeleid kunnen hieraan bezwaren verbonden zijn – voornamelijk het bezwaar dat buitenstaanders het niet meer begrijpen. Als de wiskundige expert met zeven vellen algebra bij de directeur komt en verklaart dat zijn berekeningen aantonen dat er beter 4 dan 3 reparateurs kunnen zijn, is er alle kans dat de directeur hem niet zal geloven. 'Want de helft van de tijd doen de drie reparateurs toch al niets, en daar verandert Uw abacadabra niets aan.' Het is, bij zoveel wanbegrip, verloren moeite en psychologisch onjuist om te proberen zeven pagina's algebra te gaan uitleggen. Nu dan de simulatie-aanpak.

Stel U voor een margarinefabriek met 20 pakmachines. Dit zijn volautomatische machines die een hompje margarine van het juiste gewicht (250 gram) keurig inpakken in lucifersdoosjesvorm. Wegen, vormen, inpakken en doorstromen, het gaat alles automatisch, met een tempo van 1 pakje per seconde, 60 per minuut. En dat voor 20 machines, $20 \times 60 = 1200$ pakjes per minuut. Dag en nacht, 24 uur.

De machines kunnen defect raken. Ze kunnen onnauwkeurig gaan wegen, een inpakpapiertje kan scheef komen te liggen en de regelmaat verstoren, het doorstroommechanisme kan weigeren, of de motor kan falen. Men kan werkelijk niet voorspellen

óf en wanneer iets dergelijks zal gebeuren, maar een lange ervaring heeft geleerd dat er een kans van 1 op 10 is dat een machine gedurende een bepaald uur kapot zal gaan. In een dergelijk geval moet er een reparateur bij komen. Deze kan de machine in een bepaalde tijd herstellen, afhankelijk van wat er mis is gegaan – de weeginstallatie, het doorschuifmechanisme, de motor of iets anders. De volgende gegevens vormen (naast de zojuist genoemde kans van 1 op 10) óók een deel van het model: als een machine defect raakt, dan is er een kans van

0,4 dat de reparatie $^1/_2$ uur zal vergen;
0,3 dat de reparatie 1 uur zal vergen;
0,2 dat de reparatie $1^1/_2$ uur zal vergen;
0,1 dat de reparatie 2 uur zal vergen.

Het *model* bestaande uit de kans dat een machine stuk gaat, en de dan benodigde reparatietijd, is hiermee gegeven. Het *mechanisme* om dit model te simuleren kan weer bestaan uit een zak met 100 golfballen; of laten we zeggen 100 nummers: 00, 01, 02, 03, . . . , 97, 98, 99. De procedure is als volgt:

(1) Trek een nummer uit de zak. Dit betreft het al of niet defect raken van de *eerste* machine gedurende het *eerste* uur.

(2) Indien het getrokken nummer één van de 10 nummers 90 t/m 99 blijkt te zijn, dan betekent dit dat de eerste machine gedurende dat uur defect raakt. Dit correspondeert met een kans van 1 op 10, want er zijn 100 nummers in totaal. Is het getrokken nummer 89 of lager, dan is er niets aan de hand.

(3) Indien de machine defect blijkt (dus indien één van de getallen 90 t/m 99 is getrokken), dan wordt de benodigde reparatietijd gevonden met behulp van de volgende regel: wanneer getrokken zijn de nummers

90, 91, 92 of 93, dan duurt de reparatie $^1/_2$ uur;
94, 95 of 96, dan duurt de reparatie 1 uur;
97 of 98, dan duurt de reparatie $1^1/_2$ uur;
99, dan duurt de reparatie 2 uur.

Men ziet onmiddellijk, dat de laatste regel gebaseerd is op de verhouding 4 : 3 : 2 : 1, geheel in overeenstemming met de kansen van de verschillende reparatietijden.

Deze procedure moet voor het eerste uur 20 keer worden uit-

gevoerd, want er zijn 20 machines. Het tweede nummer dat men trekt (na het eerste te hebben teruggelegd; steeds uit de volle zak met 100 nummers trekken!) betreft het al of niet stuk gaan van de *tweede* machine gedurende het *eerste* uur. Enzovoorts. Als men dit niet voor één uur, maar voor bijv. 30 uur wil herhalen, moet men op deze wijze $20 \times 30 = 600$ keer een nummertje trekken. Dit is nogal vermoeiend en wellicht zult U zich afvragen: Kan ik het zelf niet door de nummers 00, 01, . . ., 99 in willekeurige volgorde neer te schrijven in plaats van die ingewikkelde jute zak? Welnu, wij willen niet onhoffelijk lijken, maar de eerlijkheid gebiedt ons in twijfel te trekken dat U het werkelijk kunt. Wanneer een proefpersoon dergelijke getallen in willekeurige volgorde tracht op te schrijven, zal men bijv. vaak zien dat hij uit vrees te regelmatig te zijn bepaalde volgordes vermijdt die regelmatig lijken. Maar dergelijke volgordes komen nu eenmaal ook voor en wat men dan ziet is dat de proefpersoon ze met te geringe frequentie opvoert.

Er zijn twee meer bevredigende mogelijkheden, een voor thuis en een voor op de zaak. Thuis kunt U gebruik maken van tabellen van *toevalsgetallen*. Neemt men bijv. *A Million Random Digits*, een publicatie van de Amerikaanse RAND Corporation, dan vindt men op elke bladzij meer dan 1000 van dergelijke getallen vanaf 00 tot 99; en het boek bevat honderden bladzijden. De andere mogelijkheid is die van een electronische rekenmachine. Daarvan hebben wij gebruik gemaakt voor ons probleem (met behulp van de GAMMA ET van het Econometrisch Instituut). Het machineprogramma voor het trekken van toevalsgetallen is bekend en het laten draaien van het programma en laten afdrukken van de resultaten gaat razend snel. Voor 5000 getallen van 2 cijfers niet langer dan 3 minuten. Een gedeelte van deze getallen is hieronder gereproduceerd:

Machine	Uur 1	Uur 2	Uur 3	Uur 4	Uur 5	Uur 6	Uur 7
1	04	61	29	28	87	21	93*
2	87	44	07	57	65	42	57
3	12	52	79	30	25	56	35
4	98*	58	50	86	71	03	47
5	30	09	80	28	75	73	24
6	91*	07	39	93*	90*	64	96*

Machine	Uur 1	Uur 2	Uur 3	Uur 4	Uur 5	Uur 6	Uur 7
7	66	45	54	31	58	58	84
8	48	62	71	50	47	73	24
9	52	46	70	19	54	37	62
10	24	40	78	91*	33	88	39
11	49	87	79	93*	04	33	76
12	20	98*	34	43	03	50	94*
13	29	05	87	34	76	74	63
14	83	94*	03	54	06	69	06
15	81	79	40	76	47	91*	55
16	86	38	79	79	56	52	33
17	54	99*	44	15	51	07	03
18	51	14	58	31	56	73	92*
19	08	76	89	60	52	88	74
20	42	99*	48	23	34	25	52

Uit deze tabel blijkt dat er voor het eerste uur 2 getallen van 90 en hoger zijn getrokken; nl. 98 en 91 (met een sterretje aangegeven). Er zijn dus twee machines in dat uur defect geraakt. Daarvan vergt er één (98) $1^1/_2$ uur reparatietijd, de ander (91) $^1/_2$ uur. In het tweede uur zijn er moeilijkheden met vier machines (98, 94, 99, 99) en voor drie van de vier zijn ze kennelijk van ernstige aard; nl. $1^1/_2$ uur reparatietijd voor 98, 2 uur voor het tweetal van 99. In het derde uur is alles rustig: geen defecte machines. Enzovoorts. In totaal raken er 15 machines gedurende de zeven uren defect.

Wellicht is de lezer inmiddels vergeten, dat het gaat om de vraag: 3 of 4 reparateurs. Daar gaat het nog steeds om. Het is daarom van belang te weten hoeveel machines er tegelijkertijd defect zijn en onze tot op uren afgeronde tijdschaal is daarvoor te grof. Wij gaan nu met minuten werken en ook dat kan bijzonder eenvoudig. Voor elke defecte machine gaan we weer een toevalsgetal (00, 01, . . ., 99) trekken, dat het aantal minuten weergeeft waarna, in het gegeven uur, de machine defect raakt. Is het eerste toevalsgetal bijv. 17, dan betekent dit dat de eerste machine 17 minuten na het begin van het eerste uur defect raakt. Dat er niet meer dan 60 minuten in een uur zitten is geen probleem; we laten eenvoudig alle toevalsgetallen van 60 en hoger buiten beschouwing en gaan dan over naar het volgende toevalsgetal! Wij vinden:

20	24	58	(66)	(81)	07	53	(70)	47
38	26	53	(99)	01	35	(99)	12	(89)
39	(84)	51	36	28	(91)	44	15	04

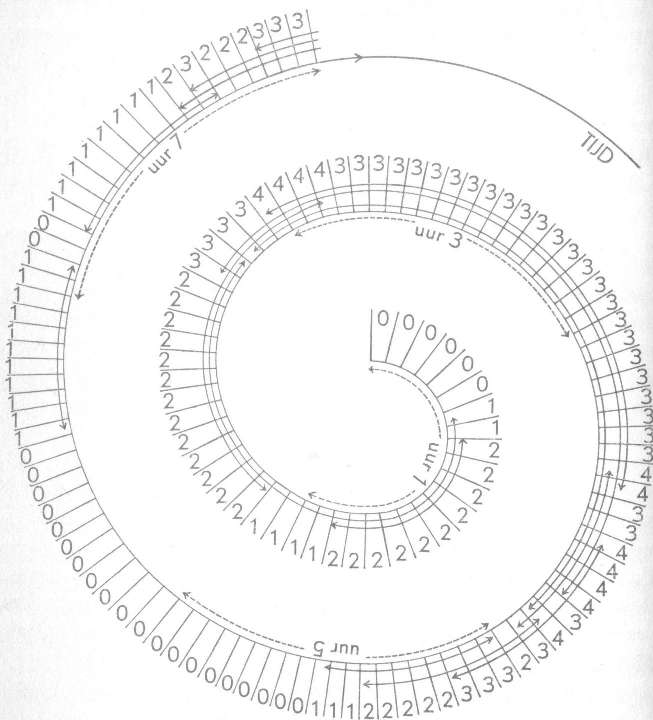
Dus raken de twee machines van het eerste uur na 20 resp. 24 minuten defect; en, zoals we gezien hebben, het kost $1^1/_2$ uur om de machine die op 0.20 uitvalt te repareren (laat 0.20 de verkorte notatie zijn voor 20 minuten na het begin van het eerste uur) en $^1/_2$ uur voor de machine die op 0.24 uitvalt. In het volgende uur raken er vier defect, nl. op 1.58, 1.07, 1.53 en 1.47; de respectieve reparatietijden zijn $1^1/_2$, 1, 2, en 2 uur. Enzovoorts.

We hebben nu alle ingrediënten bij elkaar. Wat nog gedaan moet worden is het afzetten van de resultaten op een tijdas, zodat we met één oogopslag kunnen zien hoeveel machines gelijktijdig defect zijn. Dat is gedaan op blz. 254, waar de tijdas gekromd is getekend teneinde alle zeven uren er op te krijgen. Zo zien we, dat na 20 minuten de eerste machine uitvalt en dat het $1^1/_2$ uur duurt voor hij weer werkt; dat is aangegeven door een lijn met pijlen boven de tijdas vanaf 0.20 tot 1.50. Enzovoorts. Ieder uur is ingedeeld in 20 perioden van elk 3 minuten. Voor elk van die perioden kan men onmiddellijk aan de hand van het aantal lijnen boven de tijdas nagaan, hoeveel machines er defect zijn. Dit zijn de getallen 0, 1, 2, 3, 4 in de figuur. In totaal zijn er $7 \times 20 = 140$ van dergelijke perioden in de zeven uren. Uit de grafiek blijkt, dat er 28 perioden waren waarin geen machine defect was, dus 20% van de tijd. Het volledige resultaat vindt men in tabelvorm:

Aantal defecte machines	Aantal perioden van 3 minuten	Percentage van de tijd	In uren per jaar van 360 dagen
0	28	20	1728
1	28	20	1728
2	33	24	2037
3	40	30	2469
4	11	6	679
5	0	0	0
Totaal:	140	100	8640

Natuurlijk zijn deze uitkomsten niet bijzonder nauwkeurig, maar dat doet aan het principe niets af. Men kan een langere

periode nemen, dus meer toevalsgetallen; dan zal blijken dat het ook wel eens voorkomt, dat 5 of zelfs 6 machines tegelijkertijd defect raken.



Tenslotte: het aantal reparateurs. Gewerkt wordt met een drieploegenstelsel en momenteel bestaat elke ploeg uit drie man. Moet er een vierde man bij? Welnu, zonder die vierde man gaan 679 machine-uren per jaar verloren, dus 226 voor elke ploeg. Elk verloren machine-uur kost 50 gulden. Een reparateur kost inclusief sociale lasten en wat er verder bij komt 7200 gulden per jaar. Maar $226 \times 50 = 11300$ is aanzienlijk meer, dus die vierde man moet erbij! Weliswaar valt er 20% van de tijd voor geen van hen iets te repareren, maar toch, de margarinefabriek vaart er wel bij.

Men moet zich niet laten verleiden te denken dat het bovenstaande eenvoudig was alleen omdat het probleem relatief eenvoudig was. Dit is in de wiskunde over het algemeen juist, maar bij simulatie nauwelijks. Zoals men ongeveer even goed met een roeiboot als met een vliegdekschip in een waterbak kan experimenteren, zo kan ook hier het model veel ingewikkelder worden zonder dat het mechanisme of de interpretatie van de resultaten veel moeilijker wordt. Bij wijze van voorbeeld gaan we in het bovenstaande model enige complicaties aanbrengen.

(1) De oplettende, doch wat achterdochtige lezer zal een foutje in het model hebben ontdekt: een machine die al defect is kan niet opnieuw defect raken voor hij hersteld is. Machine 20 is volgens het schema van toevalsgetallen gedurende het hele derde uur defect. Toch trokken wij daar een getal voor in het derde uur; dat bleek 48 te zijn – maar stel dat dit 91 zou zijn geweest? Hiermee kan rekening gehouden worden door als bijv. een zekere machine gedurende de eerste 7 minuten van het derde uur defect is, voor dat derde uur die machine alleen defect te laten raken als dat ná de zevende minuut gebeurt.

(2) De verdeling van de reparatieduur zal in de regel veel ingewikkelder zijn. We gingen ervan uit, dat er vier mogelijkheden zijn: $^1/_2$, 1, $1^1/_2$ en 2 uren; het kost echter weinig moeite om de simulatie-aanpak in dit opzicht realistischer te maken.

(3) Niet alle machines zullen gelijke kans hebben om defect te raken; er zijn oude en nieuwe machines!

Al deze punten zijn bij een wiskundige aanpak uiterst storend. Maar al deze complicaties worden op eenvoudige wijze in het mechanisme opgenomen door toevalsgetallen met meer cijfers te nemen. Laten we als voorbeeld een willekeurige machine in zegge het derde uur nemen, die gedurende de eerste 7 minuten van dat uur defect is. De verdere gegevens zijn als volgt.

De machine heeft per uur een kans van 0,14 om defect te raken. Dit wordt aangeduid door de eerste twee toevalscijfers: bij 00, 01, . . ., 85 gebeurt er niets, bij 86, 87, . . ., 99 raakt hij defect. Het is duidelijk, dat we op deze manier met verschillende kansen voor verschillende machines kunnen werken.

Het derde en vierde toevalscijfer geeft de reparatieduur aan

als de machine defect raakt. Dit gebeurt volgens de volgende tabel:

Reparatieduur (in minuten)	Kans	Toevalsgetal
10	0,01	00
20	0,03	01 t/m 03
30	0,05	04 t/m 08
40	0,07	09 t/m 15
50	0,09	16 t/m 24
60	0,11	25 t/m 35
70	0,13	36 t/m 48
80	0,13	49 t/m 61
90	0,11	62 t/m 72
100	0,09	73 t/m 81
110	0,07	82 t/m 88
120	0,05	89 t/m 93
130	0,03	94 t/m 96
140	0,01	97
150	0,01	98
160	0,01	99

U ziet, het is met eenvoudige middelen mogelijk een aanmerkelijk meer gedetailleerde verdeling van de reparatieduur te geven. Die verdeling kan best verschillend zijn voor verschillende machines.

Tenslotte het vijfde en zesde toevalscijfer. Die geven aan hoe laat de machine in het derde uur defect raakt *als* hij defect raakt. We delen het uur op in 100 perioden van 36 seconden elk; dus als de twee cijfers 2 resp. 5 zijn, dan betreft dit de 25ste periode van 36 seconden sinds het begin van het derde uur. Merk op dat we nog rekening moeten houden met het feit dat de machine gedurende de eerste zeven minuten van dat uur nog defect is, dus niet weer defect kan raken! Dit kan heel eenvoudig. Die 7 minuten komen (bij benadering) neer op 12 perioden van 36 seconden. Dus, als de laatste twee toevalscijfers neerkomen op 00, 01, . . ., 11, dan vervalt de mededeling van een defect.

Enkele voorbeelden:

077631: de machine raakt niet defect, want 07 is een van de getallen 00, 01, . . ., 85.

937910: de eerste twee cijfers duiden erop, dat de machine defect raakt; maar dit wordt weersproken door de laatste twee (10), die impliceren dat het defect in de eerste 7 minuten zou plaats hebben. Het defect telt dus niet.

898876: de machine raakt defect (89), dit wordt niet weersproken (76), en het kost 110 minuten om hem te repareren (88).

Zo kan men doorgaan: door toevalsgetallen bestaande uit nog meer cijfers te nemen kan men met nog meer complicaties rekening houden. Het beginsel blijft even eenvoudig. Zijn de toevalsgetallen er eenmaal, dan wordt het een zaak van nauwkeurig administreren en interpreteren, misschien een enkel tekeningetje, een beetje tellen en tenslotte een eenvoudige rekensom. Alleen blijft nog dit probleem: Hoever moet men doorgaan met simuleren? Is 7 uur voldoende? Of 70 uur? Of 700 uur? Of is 7000 uur nog niet voldoende om vertrouwen te kunnen hebben in het resultaat? Deze vraag laat zich *pragmatisch* eenvoudig beantwoorden. In dit geval is het enige waar het ons om ging het vaststellen van het percentage van de tijd dat 0, 1, 2, 3, . . . machines gelijktijdig defect zijn. Men kan nu eerst beginnen met 3 uur te simuleren, en deze percentages bepalen. Daarna 6 uur, en wèèr de percentages bepalen. Vervolgens 9 uur, en opnieuw gaat men de percentages van de tijd bepalen gedurende welke 0, 1, 2, 3, . . . machines defect zijn. Na verloop van tijd zullen dan, volgens een stelling uit de waarschijnlijkheidsrekening (de wet van de grote aantallen), de percentages die men berekent door steeds 3 uur meer te gaan simuleren niet meer merkbaar of noemenswaard van elkaar verschillen. Men kan dan met een gerust hart ophouden.

4. SCHAAK EN MONOPOLY

Het is bijna onvermijdelijk dat simulatie ook in andere omstandigheden kan worden toegepast dan juist de voorbeelden die wij noemden: de voetbalpool, de waterbak, het machinepark. Tenslotte kan men vele zaken simuleren en nabootsen, van ziekten tot bankbiljetten. Toch zou men misschien niet direct denken aan de volgende toepassing van de simulatietechniek, die we zullen inleiden aan de hand van het schaakspel en monopoly.

Met een beetje fantasie kan men het schaakspel opvatten als een miniatuur-oorlog. De symboliek is dezelfde. Alleen, als men het schaakspel met een oorlog vergelijkt, zal men wel moeten toegeven dat het realiteitsgehalte betrekkelijk gering is. Zelfs in Homerische tijden werd een oorlog niet volgens zo strikte regels op een zo beperkt gebied gevoerd. En een moderne oorlog heeft meer met roulette gemeen dan met schaak: men heeft een uiterst geringe kans om werkelijk te winnen. Daarom is het misschien beter schaak te beschouwen als een voorloper van de 'oorlogsspelen', die in later tijden in grote getale zijn geconstrueerd, en waarvan het realiteitsgehalte veel hoger is. Van dit soort oorlogsspelen, realistisch maar vredelievender dan Indiaantje spelen of zelfs een fel bevochten voetbalwedstrijd, hebt U misschien nog nooit gehoord. Een beetje pacifist zijn we allemaal, en de gedachte is beangstigend. Toch hebben ze wel degelijk hun nut.

Immers, men kan niet om der wille van de oefening een échte oorlog gaan voeren. Het mag geen 'menens' worden. Met scherp wordt niet geschoten. En desondanks moet men op de militaire academie beroepsofficieren opleiden, die, als het onverhoopt nodig zou zijn, leiding moeten kunnen geven aan een oorlog: de verbindingswegen, de fouragering, enz., en verder: wie wat waar hoe moet aanvallen. Men bootst dit nu na, *simuleert* dit, door ingewikkelde spelen met tanks, vliegtuigen, onderzeeërs, landkaarten (wegen, bruggen), spionnen, bombardementen en alles. Soms gebaseerd op historische gegevens en studies, soms ook puur verzonnen. Op deze wijze doceert men de toekomstige officieren de kunst van het tegenwoordig zo gecompliceerde krijgsmetier. Op deze wijze ook gaat men na aan welke eisen de straalvliegtuigen, de radar en het luchtafweergeschut e.d. moeten voldoen. En een hele reeks andere problemen.

Zoals schaak de voorloper is van het oorlogsspel, zo kan men monopoly de voorloper noemen van het beleidsspel. Het heeft merkwaardig lang geduurd – tot ná 1955, terwijl het eerste ingewikkelde oorlogsspel in Sleeswijk al voor 1800 was ontwikkeld –, voordat nabootsing in de vorm van spelen ook in de bedrijfswetenschappen werd toegepast. Merkwaardig, omdat de situatie in het bedrijfsleven, met concurrentie 'op leven en dood' en investeringsbeslissingen waarbij miljoenen guldens op het spel staan, frappante overeenkomsten vertoont met de krijgs-

kunde. Bovendien, ook bij de opleiding van toekomstige managers door scholen en universiteiten heeft men de moeilijkheid dat het niet goed mogelijk is de studenten een ècht bedrijf te laten leiden. Men moet dus weer op het droge leren zwemmen, en dat gebeurt met behulp van beleidsspelen (of bedrijfs- of beslissingsspelen), een vertaling van het Amerikaanse 'management games' of 'business games'. Het is een gecompliceerd monopoly. Een standaardvoorbeeld zullen we hieronder bespreken.

5. BROMFIETSEN

De problemen waarvoor de directie van een industrieel bedrijf zich geplaatst ziet vallen ruwweg uiteen in twee delen: productieproblemen en afzetproblemen. De productie vereist dikwijls een zorgvuldige organisatie, waarbij men vér vooruit moet denken, omdat vertragingen een rol van betekenis spelen. Of er nu auto's, kunstmest of boeken worden geproduceerd, men moet voor de productie van vandaag de grondstoffen gisteren hebben gekregen, en deze heeft men eergisteren moeten bestellen – of misschien wel maanden geleden, als er lange levertijden zijn, zoals met manuscripten van boeken. Bovendien moet men proberen regelmatig te produceren; de productie kan niet iedere variatie in de vraag volgen. Dit schept voorraadproblemen. Als de voorraden te klein zijn, kan men wellicht niet aan de vraag voldoen. Zijn ze te groot, dan is dit kostbaar: verzekeren, bewaken, renteverlies.

De afzet (verkoop) vormt de tweede categorie van problemen. De voornaamste instrumenten van de verkoopafdeling zijn de prijs, de reclame, afzetkanalen en service, de afbetalingsregeling en tot op zekere hoogte de kwaliteit. Tussen verkoopmanager en productiechef moet natuurlijk geregeld overleg plaats vinden. De verkoopmanager moet de productiechef vertellen naar welke kwaliteit, kleur, soort of samenstelling van het product veel vraag is. Omgekeerd zal de productiechef de verkoopmanager moeten vertellen wat technisch mogelijk, en wat technisch onmogelijk of erg kostbaar is. Gezamenlijk zullen zij er naar streven winst te maken. Uit de opbrengst van de verkoop moeten de kosten van de productie en de verkoopinspanning kunnen worden bestreden. Dit wordt alles gesimuleerd in een spel.

Het prototype van een bedrijfsspel stelt dat U directielid bent van een bedrijf op een *oligopolistische* markt; dat is een markt die door een gering aantal producenten wordt beheerst. Bijvoorbeeld de bromfietsenmarkt in Nederland. Een dergelijke oligopolistische markt laat enerzijds ruimte voor een zelfstandige verkooppolitiek (binnen zekere grenzen kunt U zelf Uw prijs vaststellen) doch anderzijds worden op zo'n markt Uw resultaten beïnvloed door de beslissingen van Uw concurrenten. Derhalve, een beetje zelfstandig, een beetje afhankelijk. Uw prijs zal door de consument, de bromfietskoper, worden vergeleken met die van Uw concurrenten. De kwaliteit zal worden vergeleken, de afbetalingsregeling, de service, en zo meer.

Als directielid bent U natuurlijk op de hoogte van een paar belangrijke aspecten van de bromfietsenmarkt. Deze worden U overigens ook in de spelregels gegeven. Daar staat dat er vier bromfietsproducenten zijn, die in het vorige jaar elk 25% van de markt beheersten – hetgeen erop neerkwam dat alle bedrijven 10.000 brommers verkochten. Zowel de productiecapaciteit (de fabriek) alsook de verkoopinspanningen van ieder der bedrijven waren het afgelopen jaar vergelijkbaar. De fabriek is in staat 1000 bromfietsen per maand te produceren, als er op volle capaciteit wordt gewerkt. De vaste kosten per jaar zijn gegeven, en de variabele kosten per fiets ook. Deze variabele kosten zullen in de regel afhangen van de productieomvang. Zij kunnen bijvoorbeeld stijgen als de capaciteit van de fabriek te dicht benaderd wordt.

Voorts is in de spelregels vermeld dat er nog 1000 fietsen in voorraad zijn aan het begin van het jaar; en dat men een stijgende trend verwacht. Maar er is een sterk seizoenspatroon: in april worden gewoonlijk vier maal zoveel bromfietsen verkocht als in november. Desgewenst kan men investeren – als men het geld heeft, of kan krijgen – in een uitbreiding van de fabriek tot een capaciteit van 1500 bromfietsen per maand; dit zal zowel de vaste als de variabele kosten beïnvloeden. De kosten van het aanhouden van voorraden zijn gegeven, bijv. 10 gulden voor iedere fiets aan het einde van de maand in voorraad. De prijs van een bromfiets bedroeg in het verleden voor elk der bedrijven 500 gulden. Verdere gegevens worden in de spelregels verschaft betreffende de afbetalingsregelingen, de mogelijke afzetkanalen, advertentiemogelijkheden en researchmogelijkheden. Er kan een mogelijkheid zijn dat U, als

U veel geld pompt in de researchafdeling van Uw bedrijf, de echt geluidloze motor kunt fabriceren en ook de meest irriterende knettermotor, zodat U in staat zult zijn aan de verlangens van de meest uiteenlopende leeftijdsgroepen te voldoen. In gecompliceerde gevallen kan de markt zijn verdeeld in deelmarkten – het Westen, het Noorden en Oosten, het Zuiden –, waar U de verkoopinstrumenten, bijv. de advertenties, verschillend kunt hanteren. Zo is er een hele reeks gegevens. Uw opdracht is om, samen met Uw drie collegadirectieleden, de fabriek gedurende de komende jaren te beheren, en daarbij winst te maken, op het marktaandeel te letten, en de fabriek in goede staat achter te laten. Hoe?

Het eerste probleem in een dergelijk spel is het vaststellen van een taakverdeling binnen het team. Een voor de hand liggende taakverdeling is bijv.: een directeur voor de productie, een directeur voor de verkoop, een directeur voor de administratie (boekhouden, winst vaststellen, kasvoorraad bepalen) en een president-directeur om de beslissingen te coördineren, met speciale opdracht om uitbreidingsinvesteringen en politiek op lange termijn vast te stellen.[1] Als dit is gebeurd begint het spel: men neemt beslissingen geldig voor de volgende maand, januari. Welke prijs, welke afbetalingsregeling, hoeveel en waar adverteren? Hoeveel produceren (maximaal 1000 per maand!), uitbreiden? Deze laatste beslissingen zullen in de regel met vertraging doorwerken. Productie in deze maand kan bijv. pas de volgende maand worden verkocht. En zo meer. Deze beslissingen gaan dan naar een scheidsrechter (dikwijls in de vorm van een electronische machine, soms alleen maar een man met een rekenliniaal) die ook de beslissingen van de *andere* bedrijven krijgt. Beslissingen betreffende *hun* prijzen, *hun* advertenties, *hun* kwaliteit, *hun* afbetalingsregeling. *Al* deze beslissingen beïnvloeden de verkopen van *ieder* bedrijf. Om deze reden is bij een enigszins ingewikkelde samenhang een rekenmachine nodig. Deze geeft binnen een enkele minuut de verkopen van ieder der bedrijven (op ieder der deelmarkten, als die

1. Slechts zelden gebeurt dit in de praktijk zo weloverwogen met titel en al. Vaak komt het voor dat iedereen zich met alles bemoeit, hetgeen leidt tot verwarring en ruzie. Soms is er – psychologen kennen dit verschijnsel goed – een 'geboren' leider, die al of niet stilzwijgend door ieder als zodanig wordt erkend. Gaat het echter slecht met het bedrijf, dan kan zo'n man zijn prestige snel verliezen.

er zijn). De verkoopresultaten van het bedrijf worden dan aan de directieleden van dat bedrijf meegedeeld. Ook andere gegevens worden verschaft, zoals de prijs die ieder der concurrenten heeft vastgesteld, en wellicht ook hun advertentiebudget. Niet vermeld wordt de interne bedrijfspolitiek van de concurrenten, zoals hun productieomvang en researchbeleid. Een aantal gegevens kan men soms als informatie bijkopen; het is dan alsof men voor een marktanalytisch onderzoek moet betalen. De gegevens behoeven natuurlijk niet alle specifiek te zijn; soms geven zij een marge aan; soms het totaal voor alle concurrenten, bijv. de concurrenten hebben in de afgelopen maand gezamenlijk in totaal 60.000 gulden aan advertenties uitgegeven.

De directieleden krijgen vervolgens tijd deze gegevens te analyseren, er conclusies uit te trekken en beslissingen te nemen voor de nieuwe periode, de maand februari. Dit herhaalt zich een aantal malen. Ieder half uur wordt op deze wijze een hele maand bedrijfsleven gesimuleerd. En de directieleden worden geconfronteerd met een hele scala van problemen die zich ook in de praktijk geregeld voordoen, zij het in een zekere gestileerde vorm, en niet in alle opzichten realistisch. Volledige realiteit is echter onbereikbaar, en er wordt ook niet naar gestreefd. Het oogmerk van dit soort spelen is dat de spelers er wat van leren. Daarom moeten er een paar problemen in zitten, waarvan wij er nu enkele nader zullen bezien.

6. PROBLEMEN BIJ BELEIDSSPELEN

Er zijn interne bedrijfsproblemen, samenhangend met het verlangen zo efficiënt mogelijk te produceren, dus de kostprijs laag te houden. Er zijn problemen op de markt, men moet het geproduceerde met winst proberen te verkopen ondanks de concurrentie. Productie en verkopen moeten zoveel mogelijk op elkaar zijn afgestemd. En er zijn problemen van bedrijfspolitieke aard: Uitbreiden? Research plegen? De administratieve problemen in een dergelijk spel komen meer naar voren bij de constructie van het spel. Ter wille van de uniformiteit van de administratie en de vergelijkbaarheid van de resultaten wordt in de regel gewerkt met door de spelleiding verstrekte formulieren; het invullen is gewoonlijk een kwestie van optellen en

aftrekken, een enkele vermenigvuldiging misschien. Dit geldt *niet* voor het maken van interessante statistieken betreffende de ontwikkeling van het marktpercentage of de voorraadomvang in de loop van de tijd; dat is een tweede taak voor de administratieve directeur.

In de meeste spelen is een frictie ingebouwd tussen productie enerzijds en verkoop anderzijds. De ene helft van het jaar overtreft de vraag de capaciteit, de andere helft blijft de vraag er ruim onder. Productieomvang en verkoopbeleid bepalen dan, tezamen met deze seizoenfluctuaties, het voorraadverloop. Er is hier een grote mogelijkheid voor een falend beleid. Dit komt dan, omdat productie en verkoop niet voldoende op elkaar zijn afgestemd en het seizoenpatroon te weinig wordt onderkend. Geregeld komt het voor dat er tegen de klippen op wordt geadverteerd, hoewel men aan de vraag niet kan voldoen! Nog veel vaker gebeurt het dat, ondanks het feit dat de voorraden groot zijn en nog groter worden, de productie gehandhaafd blijft op het punt waar de som van vaste en variabele kosten per bromfiets het laagst is. De productiechef vergeet dan dat ook voorraadkosten een rol spelen. Het bedrijf zal ook financieel op zijn tellen moeten passen. Tussen de betaling voor de grondstoffen en arbeidskosten enerzijds en de betaling van de koper voor de fiets anderzijds is er een afstand van maanden. Deze maanden moet men kunnen overbruggen. Een te liberale afbetalingsregeling die de verkoopmanager propageert om de verkopen te steunen kan leiden tot faillissement of tot peperdure leningen. Dit soort resultaten kán men voorzien, maar zij worden door de directieleden bij dit soort spelen slechts zelden onderkend.

De politiek op lange termijn is een zeer belangrijke factor in het spel. In de regel is het spel zo opgebouwd dat, indien een financieel intelligent beleid wordt gevoerd, uitbreiding mogelijk en wenselijk is. Doch als men uitbreidt zonder zich voldoende rekenschap te geven van de consequenties, zal de rampspoed snel toeslaan. Een investering kost op korte termijn veel geld, dat pas na verloop van tijd mondjesmaat zal worden terugverdiend, en men moet dus voorzichtig zijn. Zo zal men in de regel, gegeven het seizoenpatroon en de vertragingen die onvermijdelijk zijn bij de bouw van een nieuwe fabriek – bijv. een vertraging van 4 perioden: bestel de fabriek nu in mei, dan zal de productie pas in september kunnen beginnen –, moeten uit-

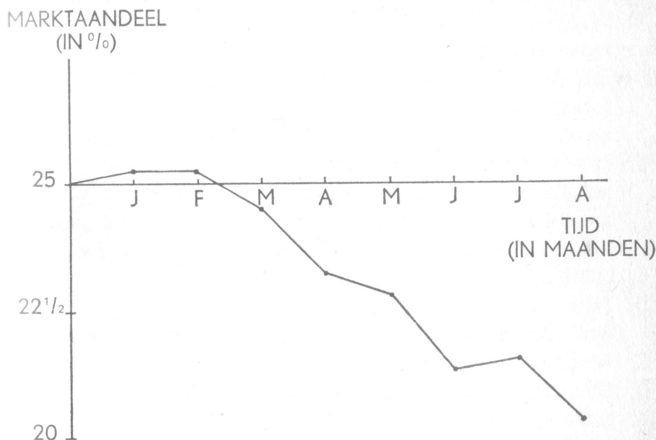
breiden in november, een maand waarin de verkopen laag zijn. De fabriek is dan klaar in maart, juist op tijd voor het hoogseizoen! In de regel zullen de spelers echter besluiten uit te breiden in juni of zo, wanneer de verkopen hoog zijn en de voorraden verdwenen. Dan zal men met een grote fabriek zitten in november, wanneer maar weinigen een bromfiets kopen. De verhoogde vaste kosten zullen dan zwaar gaan drukken. Het bedrijf is het slachtoffer geworden van gebrek aan vooruitziendheid.

Zó vermeld, lijken de problemen heel eenvoudig. In de practijk blijkt dat dit niet zo is. Het coördineren van de beslissingen, het vooruitzien, het administreren en analyseren, de vereiste snelheid en de hele atmosfeer, dat alles tezamen is voor de meeste directies te veel. Maar zelfs als men de problemen niet correct kan oplossen leert men er wel wat van. En déze vorm van leren is zo nuttig omdat de spelers zelf zo in het spel zijn opgegaan, dat ze het niet snel zullen vergeten. Uw docent kan honderd keer herhalen dat U vooruit moet denken; als U één keer failliet gaat omdat U het heeft vergeten, of althans niet heeft gedaan, blijft dat aspect van leiding geven U bij.

Afgezien hiervan zijn er nog problemen van het analyseren en interpreteren van de gegevens en de resultaten. Deze analyses zijn van belang om het U mogelijk te maken redelijke beslissingen te nemen. Zo zult u moeten trachten te bepalen hoe nauw de vraag luistert naar de prijs. Heeft een prijsverlaging van 10 gulden (dus 2%) een belangrijk effect op de vraag, vooropgesteld dat Uw concurrenten hun prijs niet wijzigen? Is het effect van het bouwen van een nieuwe winkel in Rotterdam, voor 60.000 gulden, inderdaad zes keer zo groot als het plaatsen van een advertentie van 10.000 gulden? Heeft een afbetalingsregeling die ons noopt zelf geld te lenen à 12% per jaar een voldoende effect op de vraag om deze regeling voordelig te maken – vooropgesteld dat we de gevraagde hoeveelheid kunnen leveren? Vanzelfsprekend kan men niet al deze vragen direct beantwoorden. Maar in de loop van het spel kan men er dikwijls een goed inzicht in krijgen.

Het is uiteraard zonder de spelregels volledig te vermelden en zonder enige spelperioden met beslissingen en resultaten te vermelden onmogelijk om in enig detail in te gaan op de vraag hoe de problemen moeten worden opgelost. Maar de volgende discussie in de directiekamer is wel illustratief.

Nadat een aantal perioden zijn gespeeld komt de administratieve directeur opeens met een verontrustende grafiek ter tafel: het marktpercentage is gedurende de afgelopen maanden onrustbarend gedaald:



Waar heeft het beleid gefaald? Dat vraagt de directie zich af, waarbij zich de volgende discussie ontwikkelt tussen de directeuren *A*, *B*, *C* en *D*:

A: Klaarblijkelijk zitten we met onze prijs te hoog; als puntje bij paaltje komt kijken de kopers toch alleen maar naar de prijs.

B: Maar uit de gegevens blijkt dat onze prijs eerder lager ligt dan die van de concurrenten: 490 tegen 480, 500 en 520 van de concurrenten. Aan de prijs kan 't niet liggen. Bovendien is onze marge al klein; we kunnen de prijs niet verder verlagen. Ik denk eerder dat we te weinig adverteren.

A: Maar adverteren is zo duur. En bovendien weten we van het effect van adverteren niet veel af. Daarbij komt dat we de eerste twee maanden niet meer adverteerden dan nu, en toen ging 't toch wel goed?

C: Ja, ja, dat is natuurlijk wel zo, maar hoe vaak moet ik nu nog herhalen dat alleen onze relatieve positie t.o.v. de concurrenten van belang is? Hoeveel adverteerden de anderen in 't begin, en hoeveel nu?

A: Dat weten we niet.

B: Nee, dat *weten* we niet. Maar er moet een reden zijn voor de daling in ons marktaandeel, en aan de prijzen ligt 't niet. Dan ligt het voor de hand aan adverteren te denken.

A: Waarom? Waarom? Kan het niet aan 't afbetalingssysteem liggen? Of is onze kwaliteit te laag?

D: Dat laatste kan best. We hebben de laatste tijd niets aan research gedaan. Laten we proberen de geluidloze motor te construeren. Misschien helpt ons dat uit de zorgen.

C: Is research nu wel de oplossing? In de spelregels staat dat uitgaven voor research onder de 150.000 gulden niet veel kans op succes hebben. En voor we 150.000 aan research hebben besteed... Dat wordt dus een oplossing op lange termijn.

D: Nou ja, we zullen er toch eens mee moeten beginnen. Anders blijven we zeker in 't slop.

C: Maar we hebben geen geld voor research.

D: Dan maar minder afnemerscrediet geven.

A: En nog minder verkopen?

D: Leuk is het natuurlijk niet, maar 't zal wel moeten. En dan lenen we er maar wat bij.

B: Dus het afnemerscrediet verkorten tot 2 maanden i.p.v. 4, en 30.000 aan research?

D: Liefst wat meer. 50.000? Zou dat gaan?

B: Dan moeten we lenen.

D: Goed. Verder de prijs maar handhaven. En aan advertenties zou ik ook maar niet te veel wijzigen. Wel iets minder produceren, want de voorraden zijn erg hoog. 700 fietsen lijkt me nog aan de hoge kant. We zijn over de beste tijd van het seizoen heen.

Deze discussie, tussen haakjes, is ook gesimuleerd. Maar de werkelijkheid wordt er toch wel mee benaderd. Als zodanig hoort dit gesprek in dit hoofdstuk thuis. Voor echte spelen, met spelregels en formulieren e.d., verwijzen we naar de literatuur. Enkele van de daargenoemde spelen kunnen zonder enige machinale hulp worden gespeeld.

Een verzameling opstellen over simulatie toegepast op een hele reeks van verschillende problemen kan men vinden in Guetzkow [1]. Voor de geschiedenis van beleidsspelen en een aantal voorbeelden zie Cohen and Rhenman [2]; theoretische achtergronden worden belicht door Kibbee e.a. [3]. Een verzameling beleidsspelen, van heel eenvoudig tot redelijk ingewikkeld, vindt men in het boekje van Hulshof en IJkel [4]. Dit is een geannoteerde vertaling van *Dynamic Management Decision Games* door J. R. Greene en R. L. Sisson. Het International University Contact for Management Education (kortweg IUC) beschikt over een documentatiecentrum voor beleidsspelen, dat gevestigd is in het secretariaat (Rietveldse Toorn, Delft).

[1] Guetzkow, H. (editor), *Simulation in Social Sciences: Readings*. Prentice Hall, Inc., Englewood Cliffs, N.J. 1962.

[2] Cohen, K. J., and E. Rhenman, 'The Role of Management Games in Education and Research'. *Management Science*, Vol. 7 (1961), pp. 131-166.

[3] Kibbee, J. M., C. J. Kraft and B. Nanus, *Management Games*. Reinhold Publishing Corporation, New York. 1961.

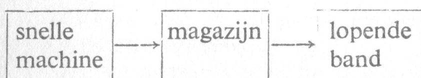
[4] Hulshof, A. H., en J. IJkel, *Beslissingsspelen*. N. Samsom N.V., Alphen aan de Rijn. 1962.

11. OVER PRODUCTIE- EN VOORRAADBESLISSINGEN

1. DE OPTIMALE SERIEGROOTTE

Wij zullen ons in dit hoofdstuk bezig houden met beslissingsproblemen op het terrein van productieomvang, voorraden en personeelsbezetting. Een belangrijk deel van deze problematiek houdt verband met de onzekerheid van het toekomstige verloop van de afzet. Hanteren we deze onzekerheid met behulp van de waarschijnlijkheidsrekening en gaan we er bovendien van uit, dat beslissingen op successieve tijdstippen dienen te worden genomen, dan komen we terecht bij de strategieën die in Hoofdstuk 7 geïntroduceerd werden. Alvorens hiertoe over te gaan zullen we het onderwerp van dit hoofdstuk met een eenvoudig voorbeeld inleiden om met een aantal begrippen kennis te maken.

In een bepaalde fabriek staat een zeer snelle machine, die per uur 200 stuks van een zeker onderdeel kan produceren. Het onderdeel wordt in een groter geheel gemonteerd op een lopende band, die een tempo heeft van 12.000 eenheden per jaar. In het algemeen zal een gereed gekomen onderdeel niet onmiddellijk op de lopende band verwerkt kunnen worden; het wordt dan eerst in een magazijn opgeslagen. Schematisch kunnen we de weg die het onderdeel aflegt dus als volgt voorstellen:



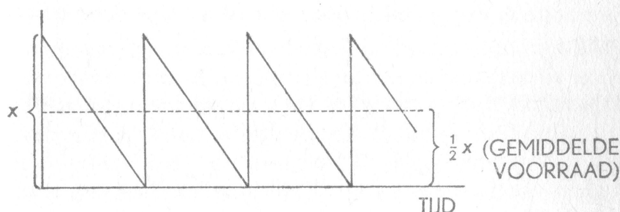
De snelle machine kan ook worden gebruikt om een aantal andere dingen te produceren. Bij het overgaan van het ene product op het andere is echter een omschakeling nodig die geld kost. Wat de omschakelingskosten betreft, is het dus voordelig om, wanneer de machine eenmaal op ons onderdeel is ingesteld, hiervan een grote serie tegelijk te maken. Anderzijds kan de lopende band de machine bij lange na niet bijhouden, zodat er steeds onderdelen opgeslagen moeten worden in het magazijn. Ook dit kost geld: afschrijving op het magazijn, onderhoud, verwarming, voorraadadministratie, rente. Dus hoe groter de voorraad, des te hoger de voorraadkosten.

Wanneer er dus bij een gegeven vraag naar het onderdeel steeds kleine series gemaakt worden, zijn de voorraadkosten laag, maar de omschakelingskosten hoog. En wanneer er grote series gemaakt worden is het omgekeerde het geval. Het ligt in de rede, dat wij gaan zoeken naar een seriegrootte waarbij de totale kosten (d.i. de som van omschakelingskosten en voorraadkosten) zo laag mogelijk zijn.

We schrijven nu a voor het benodigde aantal onderdelen per jaar en x voor het aantal dat gemaakt wordt per serie. (We nemen aan dat alle series even groot zijn.) In ons numerieke voorbeeld is $a = 12.000$; de x moeten we trachten te vinden aan de hand van ons criterium van zo gering mogelijke totale kosten. Het aantal series y dat per jaar gemaakt moet worden is gelijk aan het totaal benodigde aantal a gedeeld door de seriegrootte x:

(1) $y = a / x.$

Is er nu een bepaalde serie geproduceerd, dan gaat die naar het magazijn, waar de voorraad langzaam maar regelmatig slinkt naarmate de onderdelen door de lopende band worden afgenomen. Is de voorraad onderdelen uitgeput, dan wordt er een nieuwe serie gemaakt, en hetzelfde proces begint opnieuw. Als we in een diagram de voorraad afzetten op de verticale as en de tijd op de horizontale as krijgen we dus het volgende beeld:

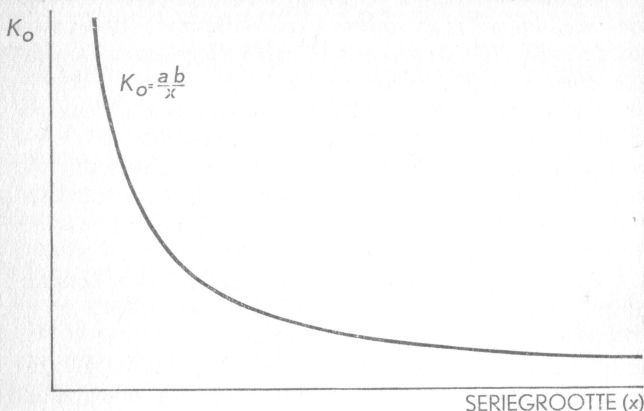


Na iedere serie beginnen we met een voorraad gelijk aan de seriegrootte x. De voorraad daalt geleidelijk (en rechtlijnig) tot hij nul is. De gemiddelde voorraad is dus $^1/_2x$.

Als de omschakelingskosten verbonden aan het produceren van een nieuwe serie b gulden bedragen, zijn de totale omschakelingskosten per jaar (K_o) gelijk aan b maal het aantal series per jaar (y):

(2) $K_o = by = ab / x$

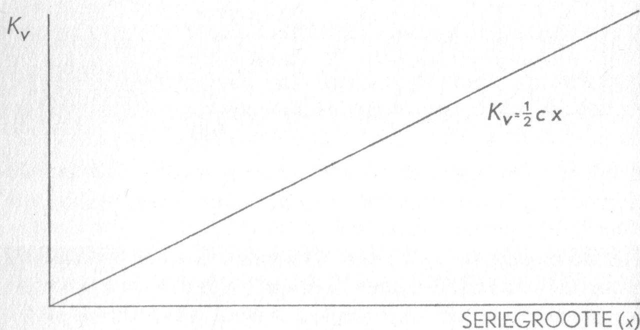
[zie formule (1)]. De totale omschakelingskosten dalen dus wanneer de seriegrootte x toeneemt; zij naderen tot nul wanneer x onbeperkt oploopt. De volgende grafiek illustreert dit.



Stel verder dat de kosten verbonden aan het opslaan van één onderdeel gedurende één jaar steeds c gulden bedragen. Gemiddeld hebben we $^1/_2x$ exemplaren in voorraad, dus de totale voorraadkosten per jaar (K_v) zijn

(3) $K_v = {}^1/_2cx.$

Grafisch kunnen we dit verband tussen voorraadkosten en seriegrootte (x) voorstellen door een rechte lijn door de oorsprong:

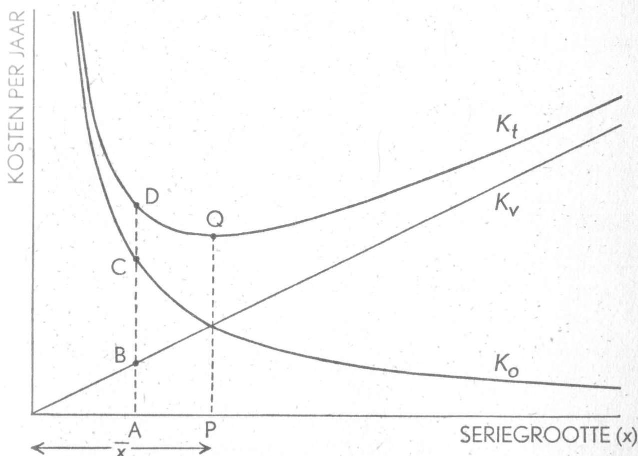


Zoals al eerder werd opgemerkt streven we er nu naar om de

som van de omschakelingskosten en de voorraadkosten zo klein mogelijk te maken. We tellen dus de formules (2) en (3) bij elkaar op. De totale kosten K_t zijn dan

$$K_t = K_v + K_o = {}^1/_2 cx + \frac{ab}{x} = \frac{{}^1/_2 cx^2 + ab}{x}$$

We kunnen nu elk van de twee afzonderlijke kostencategorieën en ook de totale kosten in één grafiek verenigen:



De curve van de totale kosten wordt gevonden door 'optelling' van de curven voor K_o en K_v. Bijvoorbeeld, wanneer de seriegrootte wordt weergegeven door het punt *A* op de horizontale as, dan zijn de voorraadkosten *AB* en de omstelkosten *AC*. Zetten we boven *AC* een lijnstuk *CD* dat even lang is als *AB*, dan stelt *AD* het bedrag van de totale kosten voor dat behoort bij de seriegrootte van punt *A*. De curve van K_t daalt in het begin (voor lage *x*-waarden), bereikt een minimum in *Q* en stijgt daarna. Uiteraard gaat het ons om dit minimum; in de eerste plaats om de seriegrootte waarvoor dit minimum wordt bereikt (de optimale seriegrootte, weergegeven door $\bar{x}$) en in de tweede plaats het met *x* verbonden totale kostenbedrag. Dit is in de grafiek aangegeven door het lijnstuk *PQ*, hetgeen dus het laagste totale kostenbedrag is dat gerealiseerd kan worden.

Teneinde de optimale seriegrootte $\bar{x}$ te bepalen zullen wij een kleine kunstgreep toepassen. We trekken nl. van de totale kosten de vierkantswortel van $2abc$ af. Het resultaat is als volgt:

$$K_t - \sqrt{2abc} = \frac{{}^1/_2 cx^2 + ab}{x} - \sqrt{2abc}$$

$$= \frac{{}^1/_2 cx^2 + ab - x\sqrt{2abc}}{x}$$

$$= \frac{(x\sqrt{{}^1/_2 c} - \sqrt{ab})^2}{x} ,$$

waarbij gebruik is gemaakt van de regel $(X - Y)^2 = X^2 + Y^2 - 2XY$ uit de elementaire algebra:

$X = x\sqrt{{}^1/_2 c}, \quad Y = \sqrt{ab},$

$X^2 = {}^1/_2 cx^2, \quad Y^2 = ab, \quad 2XY = 2x\sqrt{{}^1/_2 abc} = x\sqrt{2abc}.$

Tellen we vervolgens links en rechts $\sqrt{(2abc)}$ erbij op, dan vinden we:

$$(4) \quad K_t = \sqrt{2abc} + \frac{(x\sqrt{{}^1/_2 c} - \sqrt{ab})^2}{x} .$$

De totale kosten worden aldus beschouwd als de som van twee uitdrukkingen. De eerste is $\sqrt{(2abc)}$, die onafhankelijk is van x. In gewoon Nederlands wil dit zeggen, dat hoe we de seriegrootte ook kiezen, deze uitdrukking blijft wat hij is; hij speelt dus geen rol bij onze pogingen om de seriegrootte zodanig aan te passen, dat de totale kosten zo gering mogelijk zijn. De tweede uitdrukking bevat x wel degelijk; preciezer, hij is een breuk met een kwadraat in de teller en x in de noemer. Twee overwegingen zijn nu van belang. Ten eerste: die noemer is steeds positief. Dit volgt uit het eenvoudige feit, dat een seriegrootte nu eenmaal niet nul of negatief kan zijn. Ten tweede: de teller is hetzij positief hetzij nul. Dit volgt uit het feit, dat een kwadraat niet negatief kan zijn.

Aldus zien we, dat de totale kosten gelijk zijn aan een constant bedrag, $\sqrt{(2abc)}$, plus een breuk waarvan de teller hetzij nul hetzij positief is en de noemer steeds positief. De breuk is dus zelf hetzij nul hetzij positief. De totale kosten zijn zo gering mogelijk wanneer die breuk zo klein mogelijk is; en de laagste waarde van die breuk is nul. Dus is de minimale waarde van de totale kosten gelijk aan $\sqrt{(2abc)}$, hetgeen de lengte is van de lijn

PQ in de figuur op blz. 271. Voor welke seriegrootte wordt dit minimum bereikt? Welnu, dit minimum vereist, dat we de breuk nul maken en dus ook de teller van de breuk. Gaan we terug naar formule (4), dan zien we dat $x\sqrt{(^1/_2 c)}$ gelijk moet worden gemaakt aan $\sqrt{(ab)}$. Dit leidt tot de volgende formule voor de optimale seriegrootte:

$$\bar{x} = \frac{\sqrt{(ab)}}{\sqrt{(^1/_2 c)}} = \sqrt{\frac{2ab}{c}}.$$

Voorbeeld: we gaan uit van onze jaarproductie van 12.000 stuks ($a = 12.000$). Stel dat de kosten van omschakeling per keer f 40 bedragen ($b = 40$) en de voorraadkosten f 6 per stuk per jaar ($c = 6$). Dan is de optimale seriegrootte

$$\bar{x} = \sqrt{\frac{2 \times 12.000 \times 40}{6}} = \sqrt{160.000} = 400.$$

Onze snelle machine dient dus het onderdeel in series van 400 te maken. De gemiddelde voorraad is dan $^1/_2 \times 400 = 200$, zodat de voorraadkosten gelijk zijn aan $200 \times 6 = 1200$ gulden per jaar. Er wordt $12.000 / 400 = 30$ maal per jaar omgeschakeld; de omschakelingskosten zijn dus $30 \times 40 = 1200$ gulden per jaar, hetgeen even hoog is als de voorraadkosten.[1] De totale kosten zijn dus f 2400 per jaar. We vonden hierboven, dat het minimale bedrag van de totale kosten is

$$\sqrt{2abc} = \sqrt{2 \times 12.000 \times 40 \times 6} = 2400,$$

dus inderdaad hetzelfde bedrag.

Kunnen wij het resultaat ook intuïtief duidelijk maken? Dat kan, althans tot op zekere hoogte. Hoe hoger de kosten per omschakeling (b), des te belangrijker wordt het niet zo vaak om te schakelen en dus langere series te maken; en inderdaad laat de formule voor de optimale seriegrootte $\bar{x}$ zien, dat de series groter dienen te zijn naarmate b groter is. Wanneer daarentegen de voorraadkosten per stuk (c) groter worden, daalt blijkens dezelfde formule de optimale seriegrootte. Ook dit ligt voor de hand; want als het erg kostbaar gaat worden

1. Dit is niet de toevallige consequentie van ons getallenvoorbeeld. Het is steeds zo (gegeven de hier gemaakte veronderstellingen), hetgeen de lezer gemakkelijk zelf kan verifiëren door $\bar{x}$ in de twee afzonderlijke kostenfuncties te substitueren. Meetkundig bezien komt dit resultaat hierop neer, dat de lijn PQ van de figuur op blz. 271 gaat door het snijpunt van de twee kostenlijnen.

voorraad te houden moeten we die klein houden, en de enige manier om dit te verwerkelijken is met grote frequentie kleine series te maken. En wat de jaarproductie (*a*) betreft, de formule laat zien dat de optimale seriegrootte stijgt naarmate *a* groter is. Om dit te doorzien dienen wij ons te realiseren, dat bij een gegeven seriegrootte die voorraadkosten onafhankelijk zijn van *a* ($K_v = {}^1/_2cx$ is bij gegeven *x* alleen maar afhankelijk van *c*), dat daarentegen de omschakelingskosten evenredig zijn met *a*. Het ligt daarom in de rede, dat men bij grotere *a* zal trachten op de omschakelingskosten te bezuinigen. En dit is precies wat de formule ons vertelt; die zegt nl. dat de optimale seriegrootte bij klimmende jaarproductie stijgt, dus dat de series groter dienen te worden.

2. GEVOELIGHEIDSANALYSE

Wij hebben het vraagstuk van de optimale seriegrootte onder de gestelde voorwaarden opgelost, maar al gauw rijst dan de vraag: hoe 'gevoelig' is dit optimum? De betekenis van deze vraag komt op het volgende neer. Blijkens de formule voor de optimale seriegrootte hangt deze o.a. af van de voorraadkosten per stuk per jaar (*c*). We hebben gezien, dat deze voorraadkosten op hun beurt weer bepaald worden door een aantal factoren, zoals de afschrijving op het magazijn e.d. Hoeveel men per jaar op het magazijn dient af te schrijven is lang niet altijd een eenvoudig probleem. Het zal dus vaak zo zijn, dat het als voorraadkosten opgevoerde bedrag gekenmerkt is door een zekere foutenmarge. En deze fouten werken door in de formule voor de optimale seriegrootte. Past men die formule toe, dan zal men dikwijls tot een uitkomst komen, die verschilt van de werkelijk optimale seriegrootte. Hierop nu doelt de vraag, die aan het begin van deze alinea werd gesteld: als de gekozen seriegrootte afwijkt van de optimale, wat zijn dan de consequenties? Welnu, één ding kunnen we onmiddellijk hierop zeggen: de totale kosten zullen hoger dan het minimum zijn, zowel wanneer de gekozen seriegrootte beneden het optimum ligt als wanneer die erboven ligt. Dit volgt uit de figuur op blz. 271, die immers laat zien, dat de curve voor K_t voortdurend omhoog klimt wanneer we ons vanuit *Q* hetzij naar links hetzij naar rechts bewegen. Maar de gevoeligheidsanalyse is nog iets

preciezer in zijn vraagstelling: als de gekozen seriegrootte zus-en-zo is, met *hoeveel* stijgen dan de totale kosten boven het minimum? Met dit probleem, dat in het onderhavige geval niet moeilijk is, zullen wij ons nu bezig houden.

Wanneer de gekozen seriegrootte 10% boven het optimum ligt, betekent dit dat hij gelijk is aan $1{,}1\bar{x}$; ligt hij 30% beneden het optimum, dan is hij gelijk aan $0{,}7\bar{x}$; enz. In het algemene geval is de seriegrootte $(1 + \delta)\bar{x}$, waarbij δ staat voor de relatieve fout; dus is $\delta = 0{,}1$ wanneer de gekozen seriegrootte 10% boven het optimum ligt, $\delta = -0{,}3$ wanneer hij 30% er beneden ligt, enz. Nu wordt voor elke seriegrootte x (en dus ook voor elke waarde van δ) het bedrag van de totale kosten door formule (4) gegeven. We vullen $x = (1 + \delta)\bar{x}$ in en vinden dan:

$$(5) \quad \sqrt{2abc} + \frac{[(1 + \delta)x\sqrt{{}^1/_2 c} - \sqrt{ab}]^2}{(1 + \delta)\bar{x}},$$

hetgeen het kostenbedrag voorstelt in het geval van een relatieve fout δ t.o.v. de optimale seriegrootte. De formule ziet er wat ingewikkeld uit, maar hij kan drastisch vereenvoudigd worden. Daartoe roepen wij in herinnering, dat de optimale seriegrootte $\bar{x}$ gelijk is aan de vierkantswortel van $2ab/c$. Dus is

$$\bar{x}\sqrt{{}^1/_2 c} = \sqrt{\frac{2ab}{c}} \times \sqrt{{}^1/_2 c} = \sqrt{ab},$$

zodat de uitdrukking tussen vierkante haken in (5) wordt:

$$(1 + \delta)\bar{x}\sqrt{{}^1/_2 c} - \sqrt{ab} = (1 + \delta)\sqrt{ab} - \sqrt{ab} = \delta\sqrt{ab}.$$

De teller van de breuk in (5) is hiervan het kwadraat, dus $\delta^2 ab$. De breuk als geheel wordt dan

$$\frac{\delta^2 ab}{(1 + \delta)\sqrt{(2ab/c)}} = \frac{{}^1/_2\delta^2}{1 + \delta}\sqrt{2abc}.$$

Het totale kostenbedrag (5) vinden we tenslotte door optelling:

$$\sqrt{2abc} + \frac{{}^1/_2\delta^2}{1+\delta}\sqrt{2abc} = \sqrt{2abc}\left\{1 + \frac{{}^1/_2\delta^2}{1+\delta}\right\},$$

hetgeen een aanmerkelijke vereenvoudiging t.o.v. formule (5) is.

We kunnen nog een verdere vereenvoudiging bereiken zodra we ons herinneren, dat $\sqrt{(2abc)}$ het minimale totale kostenbedrag voorstelt (dat bereikt wordt als de optimale seriegrootte $\bar{x}$ wordt toegepast). Laten we er dan toe overgaan de kosten

in dit minimale bedrag uit te drukken. Dit betekent niets anders dan dat we het kostenbedrag (6) delen door $\sqrt{(2abc)}$. Het resultaat is

$$(7)\quad 1 + \frac{{}^1/_2\delta^2}{1+\delta},$$

en dit is (indien met 100 vermenigvuldigd) het kostenbedrag bij een fout δ, uitgedrukt als percentage van het minimale kostenbedrag. We zien, dat dit resultaat onafhankelijk is van de coefficiënten a, b en c, hetgeen de discussie sterk vereenvoudigt. Neem dan, als voorbeeld, $\delta = 0{,}1$; de gekozen seriegrootte ligt dus 10% boven het optimum. Dan wordt de uitdrukking (7)

$$1 + \frac{{}^1/_2 \times 0{,}01}{1{,}1} = 1{,}0045;$$

in procenten 100,45. Dit betekent, dat als de gekozen seriegrootte de optimale waarde met 10% overschrijdt, de resulterende extra-kosten nog geen half procent bedragen. De 'straf' verbonden aan een dergelijke overschrijding is dus bijzonder gering. Op analoge wijze kunnen we de extra-kosten berekenen wanneer δ andere waarden aanneemt. Dit is hieronder gedaan voor procentuele fouten (100δ) van 20, 30, . . . , 100%:

Fout (in %)	20	30	40	50	60	70	80	90	100
Extra-kosten (in %)	1,7	3,5	5,7	8,3	11	14	18	21	25

Het resultaat laat zien, dat de procentuele extra-kosten zeer gering zijn zolang betrekkelijk weinig van de optimale seriegrootte wordt afgeweken, maar ook dat die extra-kosten de neiging hebben snel te klimmen wanneer de afwijkingen groter en groter worden. Om een extreem geval te nemen, beschouw een gekozen seriegrootte 10 maal zo groot als de optimale. Dan is $\delta = 9$ en de uitdrukking (7) heeft tot uitkomst 5,05. De totale kosten van omschakeling en voorraad houden zijn dus ruim het vijfvoudige van het noodzakelijke minimum!

Voor afwijkingen *beneden* de optimale seriegrootte loopt de afleiding op precies dezelfde manier. We hebben dan $\delta = -0{,}1$ voor een afwijking van 10%, enz., en de resultaten zijn als volgt:

Fout (in %)	10	20	30	40	50
Extra-kosten (in %)	0,6	2,5	6,4	13	25

Ook hier zien we, dat de extra-kosten procentueel zeer gering zijn zolang de fout procentueel matig is, maar daarna snel oplopen. De uitkomsten wekken de indruk, dat fouten beneden het optimum ernstiger zijn dan overeenkomstige fouten boven het optimum, maar dat is meer schijn dan werkelijkheid. Neem bijv. het geval dat we een factor 2 mis zijn. Boven het optimum betekent dit, dat de gekozen seriegrootte $2\bar{x}$ is; beneden het optimum dat de seriegrootte $^1/_2\bar{x}$ is. Het eerste geval correspondeert met $\delta = 1$ (100%), het tweede met $\delta = -0{,}5$ (50%), en we hebben hierboven gezien dat in beide gevallen de extra-kosten 25% van het minimale totale kostenbedrag uitmaken. Overigens behoeft het bepaald niet in alle situaties zo te zijn, dat de kostenconsequenties 'symmetrisch' zijn voor afwijkingen boven en beneden het optimum.

Samenvattend kunnen we dus stellen, dat binnen het raam van de hier gemaakte veronderstellingen de formule voor de optimale seriegrootte weinig gevoelig is in de zin dat afwijkingen van het optimum tot relatief geringe extra-kosten aanleiding geven. Hetgeen betekent, dat het niet noodzakelijk is om de bepalende factoren (de coëfficiënten a, b en c) met zeer grote nauwkeurigheid te kennen. Maar er is een voorbehoud: men mag het met die afwijkingen ook weer niet te bont maken, want de kostennadelen worden zeer aanzienlijk wanneer de gekozen seriegrootte te zeer van het optimum verschilt.

3. EEN UITVOERIGER KOSTENFUNCTIE

'Binnen het raam van de hier gemaakte veronderstellingen' was de clausule, die we aan de samenvatting van de gevoeligheidsanalyse verbonden. Zijn die veronderstellingen niet realistisch, dan helpt een dergelijke analyse weinig; wat wij dan moeten doen is het opzetten van een theorie die beter aan de eisen van realisme voldoet. En hoe staat het met het realisme van de tot nu toe gebruikte veronderstellingen? Welnu, een ervan houdt in, dat de jaarproductie van het onderdeel (a) een gegeven constante is. In werkelijkheid zal men die productie aan de vraag moeten aanpassen. Is hij systematisch groter dan de vraag, dan hopen zich voorraden op; is hij systematisch lager, dan raken de voorraden op. Maar de toekomstige vraag is onbekend en dit introduceert het aspect van de onzekerheid. Wij

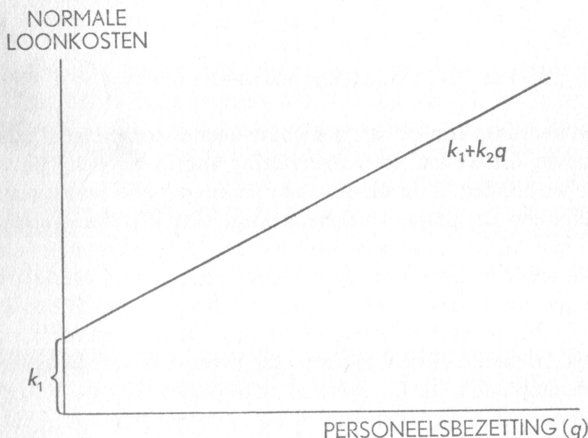
zullen nu een uitvoeriger kostenfunctie behandelen aan de hand van een studie uitgevoerd door Holt, Modigliani, Muth en Simon.

De studie heeft betrekking op de verfafdeling van de Pittsburgh Plate Glass Company. Evenals in de voorgaande bladzijden gaat het om productie tegen zo gering mogelijke kosten, maar het aantal kostencategorieën dat nu onderscheiden zal worden is een viertal:[1] *voor noot zie pagina 279*

1. *Normale loonkosten*, ruim geïnterpreteerd, dus inclusief sociale lasten. Indien de fabrieksdirecteur besluit de productie te verhogen, dan dient het aantal manuren te worden verhoogd, hetzij door overwerk hetzij door aantrekking van nieuw personeel. Overwerk zullen wij hier buiten beschouwing laten evenals de specifieke kosten verbonden aan het aantrekken als zodanig; deze kostencategorieën worden hieronder nader bezien. Het gaat hier om de kosten van de personeelsbezetting, gegeven een bepaald productieniveau dat gedurende lange tijd constant wordt verondersteld. In eerste benadering worden deze kosten als een lineaire functie van de personeelsbezetting beschouwd:

$$k_1 + k_2q,$$

waarbij q de personeelsbezetting voorstelt en k_1 en k_2 als constante coëfficiënten geïnterpreteerd worden. Grafisch:

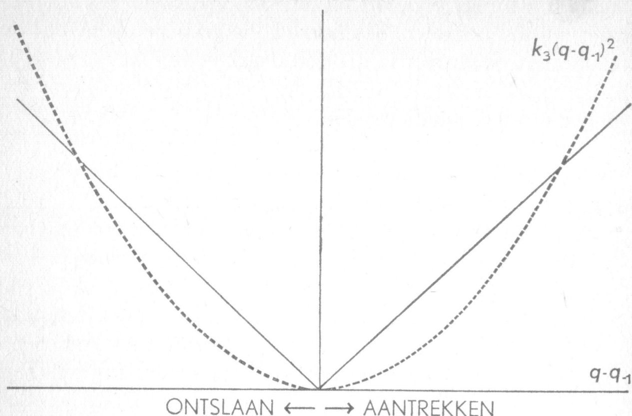


De vaste kosten van deze categorie worden gerepresenteerd

door k_1, de variabele door k_2. Uiteraard veranderen deze coefficiënten wanneer bijv. het loonniveau verhoogd wordt. Hier worden zij als gefixeerd beschouwd, hetgeen betekent dat de berekeningen, waarover in dit hoofdstuk verder wordt gesproken, herhaald dienen te worden wanneer dergelijke loonsverhogingen zouden plaats hebben.

2. *De kosten van het aantrekken en het ontslaan van personeel.* Wanneer een nieuweling wordt aangetrokken moet hij worden ingewerkt; desondanks krijgt hij i.h.a. onmiddellijk een volledig loon. Maar zelfs vóór hij wordt ingewerkt worden er kosten gemaakt: de kosten van het selecteren, van de controle van gegevens, van het medisch onderzoek, enz. Dit alles betekent, dat een vergroting van de personeelsbezetting met kosten gepaard gaat en wel des te meer naarmate die vergroting omvangrijker is. Echter, ook bij *inkrimping* van het personeel worden kosten gemaakt. Een man die ontslagen wordt krijgt een zekere som bij zijn vertrek; wanneer de inkrimping enigszins omvangrijk is zal het vaak nodig zijn zekere diensten te reorganiseren, en ook dat kost geld; de contracten met Amerikaanse vakverenigingen bevatten vaak bepalingen, die kosten impliceren bij ontslag van personeel; en het verlenen van ontslag is niet populair, hetgeen het bedrijf op prestigeverlies kan komen te staan en daarmee op moeilijkheden bij het aantrekken van personeel in betere tijden. (Uiteraard is het niet bepaald eenvoudig dit prestigeverlies in geld uit te drukken, maar meestal geldt in dergelijke gevallen, dat een ruwe schatting de voorkeur verdient boven het verwaarlozen van het verlies.) Al met al zien we dat ook inkrimping van het personeel met kosten gepaard gaat en wel des te meer naarmate de inkrimping omvangrijker is. Gaan we ervan uit, dat zowel bij uitbreiding als bij inkrimping de kosten zich lineair gedragen, dan krijgen we een kostenfunctie die de vorm heeft van de letter V. Een afbeelding hiervan kan men aantreffen op de volgende bladzijde bovenaan. Langs de horizontale as is de verandering van de personeelsbezetting afgezet, $q - q_{-1}$, waarbij q_{-1} de personeelsbezetting van de vorige periode voorstelt. (In de hier besproken studie

1. In hetgeen volgt zullen wij ons beperken tot die kostencategorieën die rechtstreeks verband houden met beslissingen inzake productie en personeelsbezetting. Voor deze categorieën zal het kostenbegrip echter ruim worden geïnterpreteerd.



wordt met perioden van een maand gewerkt.) Rechts van de verticale as is de verandering van de personeelsbezetting positief; dat is het geval van het aantrekken van nieuw personeel. Links van de verticale as is de verandering negatief; daar wordt personeel ontslagen. Men ziet gemakkelijk, dat de V-lijn impliceert dat de kosten evenredig zijn met de omvang van de verandering.

In de figuur is ook een gestippelde kromme getekend, die voor kleine veranderingen van de personeelsbezetting beneden de V-lijn ligt en voor grote veranderingen daarboven. De vergelijking van die kromme is

$$k_3(q - q_{-1})^2,$$

dus een constante (k_3) vermenigvuldigd met het kwadraat van de verandering van de personeelsbezetting. In tegenstelling tot de V-lijn houdt deze specificatie in, dat de kosten meer dan evenredig met de verandering toenemen: in het begin (bij kleine veranderingen) zijn de kosten heel gering, later – naarmate de veranderingen groter worden – stijgen zij sneller en sneller. Een dergelijke specificatie kan realistisch zijn, wanneer men bij omvangrijker uitbreidingen op voortdurend grotere hindernissen stuit; bijv. wanneer het arbeidspotentieel in de onmiddellijke omgeving van de fabriek uitgeput raakt en men genoodzaakt wordt om bussen te laten rijden om ver weg wonende arbeiders te halen en te brengen. In hetgeen volgt zullen wij

ons beperken tot de kwadratische specificatie; de V-lijn zullen we dus laten vallen.

Nog één opmerking over deze kostencategorie. Onze specificatie en ook de figuur gaan ervan uit, dat de kosten verbonden aan het aantrekken van een zeker aantal arbeiders gelijk zijn aan die van het afstoten van een gelijk aantal. Bijvoorbeeld, laat de uitbreiding 10 man bedragen, dus $q - q_{-1} = 10$; dan zijn de daaraan verbonden kosten $k_3 \times 10^2 = 100k_3$. Neem vervolgens het geval van een inkrimping van 10 man, dus $q - q_{-1} = -10$, dan zijn de kosten $k_3 \times (-10)^2 = 100k_3$, dus precies even hoog. Uiteraard is het geen Wet van Meden en Perzen dat de kosten van aantrekken en afstoten van personeel aldus 'symmetrisch' zijn; de kosten van aantrekken kunnen lager zijn, en ook hoger, dan die van het afstoten van een gelijk aantal arbeiders. Het is mogelijk de kwadratische specificatie zodanig te amenderen dat hiermee rekening wordt gehouden, maar dat is een verfijning waarop wij niet zullen ingaan.

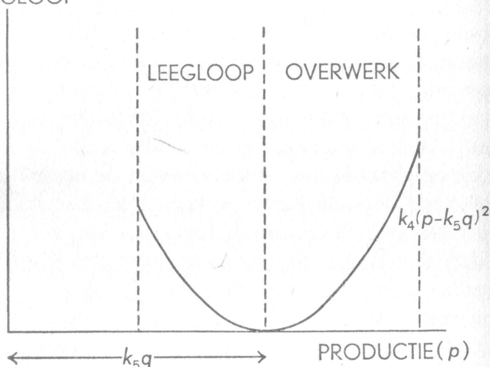
3. *Kosten van overwerk en leegloop*. Bij de bespreking van de normale loonkosten werd opgemerkt, dat een verhoogde productie gerealiseerd kan worden hetzij met personeelsuitbreiding hetzij door het aanwezige personeel overwerk te laten verrichten. De uitbreiding is in zoverre nadelig dat hiermee kosten gepaard gaan – dat hebben we zojuist gezien; het alternatief is overwerk, dat i.h.a. ongeveer 50% duurder uitkomt dan het normale werk. Waardoor worden de kosten van overwerk bepaald? Welnu, overwerk is pas nodig tegen de tijd dat de productie zo groot is dat hij niet meer in de normale tijd door het aanwezige personeel kan worden gerealiseerd. Laten we k_5 schrijven voor de maximale productie van een gemiddelde arbeider; dan is k_5q de maximale productie waartoe de huidige personeelsomvang (q) in staat is zonder overwerk. We zullen p schrijven voor de productie; kennelijk moet er dan overwerk verricht worden als p boven k_5q ligt, anders gezegd, als $p - k_5q$ positief is. En wat als $p - k_5q$ negatief is? In dat geval ligt de productie beneden het niveau dat door het aanwezige personeel kan worden gerealiseerd. Er is dan sprake van leegloop. In een aantal situaties is het denkbaar, dat men het personeel onderhoudswerkzaamheden laat verrichten, die anders achterwege zouden zijn gebleven; hoe dan ook, een deel van

het personeel is niet of minder productief, hetgeen betekent dat ook leegloop met kosten gepaard gaat.

Zo zien we, dat zowel wanneer $p - k_5q$ positief is (overwerk noodzakelijk) als wanneer $p — k_5q$ negatief is (leegloop) kosten gemaakt worden; in het ene geval kosten van overwerk, in het andere kosten van leegloop. Hoe zal de wiskundige vorm van het verband tussen kosten en $p — k_5q$ zijn? Om deze vraag te kunnen beantwoorden dienen we te weten, dat de arbeiders van de fabriek alle in zekere mate op hun terrein gespecialiseerd zijn. Een kleine productietoeneming zal daarom slechts overwerk vragen van een beperkt aantal employé's, nl. die welke in knelpunt ('bottleneck') functies werken. Bij grotere productietoenemingen is het niet alleen noodzakelijk meer employé's overwerk te doen verrichten, maar bovendien is in toenemende mate toezicht en coördinatie vereist en wel op tijden die minder gunstig zijn dan de normale arbeidsuren. De kosten stijgen dus meer dan evenredig en een kwadratische functie is dan een redelijke eerste benadering:

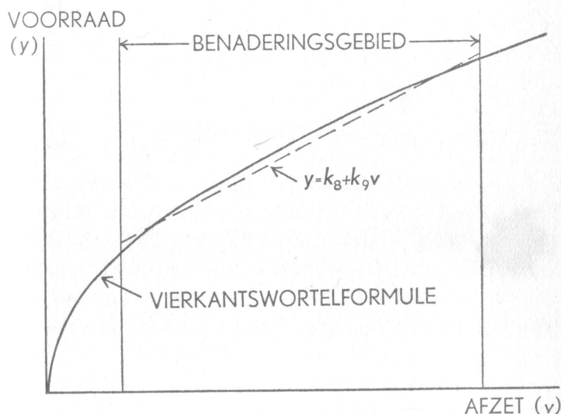
$$k_4(p — k_5q)^2.$$

In onderstaande figuur wordt dit verband grafisch toegelicht:



4. *Voorraad- en omschakelingskosten.* Tenslotte hebben we de kostencategorieën, die we al in het begin van dit hoofdstuk onder de loupe namen. Onze fabriek produceert verf op voorraad; en het houden van voorraden kost meer geld naarmate de voorraad groter is. Wij hebben eerder gezien, dat onder

zekere voorwaarden de voorraad gemiddeld genomen gelijk is aan de helft van de seriegrootte en dat de optimale seriegrootte evenredig is met de vierkantswortel van de hoeveelheid die per tijdseenheid geleverd dient te worden; dus is ook onder dezelfde voorwaarden de optimale voorraad (gemiddeld genomen) evenredig met dezelfde vierkantswortel. We hebben óók gezien, dat onder dezelfde voorwaarden de seriegrootte niet erg gevoelig is in de zin, dat tamelijk grote afwijkingen van het optimum tot slechts geringe extra-kosten aanleiding geven. Hetzelfde geldt dan voor de optimale voorraad; en we zullen van dit resultaat gebruik maken teneinde een benadering door te voeren, die voor ons doel van practische betekenis zal zijn. Zouden wij nl. het zojuist genoemde resultaat letterlijk opvatten, dan zou de optimale voorraad evenredig zijn met de vierkantswortel van de afzet van de verf; maar in feite zullen we die relatie lineair benaderen, d.w.z. we zullen de optimale voorraad schrijven als $k_8 + k_9v$, waarbij de k's coëfficiënten zijn en v de verkoop van verf voorstelt. Onderstaande figuur laat zien, dat binnen een beperkt benaderingsgebied de lineaire benadering van redelijke kwaliteit kan zijn:



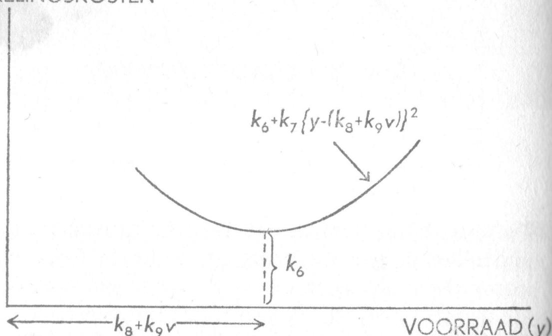
De toevoeging 'binnen een beperkt benaderingsgebied' moet uitdrukkelijk worden gemaakt, want de lineaire benadering wordt slecht wanneer we haar over te grote gebieden wensen toe te passen. Overigens pleegt de afzet van onze verffabriek in feite binnen zekere marges te fluctueren, die weliswaar bepaald

niet dicht bij elkaar liggen maar aan de andere kant ook weer niet extreem zijn.[1] Voorts moet men niet uit het oog verliezen hetgeen in het begin van dit hoofdstuk al is gesteld, nl. dat zelfs tamelijk grote afwijkingen van het optimum (veroorzaakt door de lineaire benaderingsformule) tot geringe extra-kosten aanleiding geven.

Wij zullen de voorraad- en omschakelingskosten nu schrijven als de som van een tweetal bedragen. Het eerste is het minimum dat voor deze categorie denkbaar is, te schrijven als k_6. Het tweede wordt bepaald door de afwijking tussen de feitelijke voorraad en de optimale waarde, dus door het verschil tussen y en $k_8 + k_9v$. Het is in dit verband nuttig op te merken, dat er óók kosten worden gemaakt als y kleiner is dan het optimum $k_8 + k_9v$. Wanneer nl. de voorraad te klein wordt, loopt men het risico niet te kunnen leveren wanneer de vraag plotseling stijgt; de fabriek is dan genoodzaakt de klant te laten weten, dat hij gedurende zekere tijd op de order dient te wachten. Gebeurt dit op enige schaal, dan worden de klanten ongeduldig en bestaat het risico dat zij zich tot de concurrentie wenden. Bovendien stijgen de omschakelingskosten bij te laag voorraadniveau; de apparatuur moet immers steeds opnieuw worden ingesteld op een andere verfsoort naarmate telkens een soort uitverkocht raakt. De gekozen wiskundige vorm is weer van het kwadratische type:

$$k_6 + k_7[y - (k_8 + k_9v)]^2$$

en de grafiek is de volgende: *voor noot zie pag. 285*



Duidelijk blijkt uit deze grafiek, dat k_6 de minimale waarde is van de voorraad- en omschakelingskosten, die gerealiseerd wordt als de voorraad de optimale waarde $k_8 + k_9v$ aanneemt.

Samenvattend hebben we dus een viertal kostencategorieën, die elk bepaald worden door zekere constanten (de negen k's) en door vier (maandelijkse) variabelen: productie (p), personeelsbezetting (q), voorraad (y) en verkoop (v). De totale kosten worden gevonden door optelling:

Totale kosten per maand

$= k_1 + k_2q$	(normale loonkosten)
$+ k_3(q - q_{-1})^2$	(aantrekken en afstoten)
$+ k_4(p - k_5q)^2$	(overwerk en leegloop)
$+ k_6 + k_7[y - (k_8 + k_9v)]^2$	(voorraad en omschakeling)

4. HET PROBLEEM VAN DE KOSTENMINIMERING

Bij de bespreking van de voorraad- en omschakelingskosten werd de term 'optimale voorraad' gebruikt; het komt erop neer, dat we de voorraad met de afzet laten op en neer bewegen volgens de formule $y = k_8 + k_9v$, met het gevolg dat de voorraad- en omschakelingskosten dan hun minimale waarde k_6 bereiken. Bij de kosten van overwerk en leegloop hebben we dezelfde situatie: indien onze fabrieksdirecteur zodanig te werk gaat dat p gelijk is aan k_5q (dus productie conform de capaciteit van de fabriek op basis van de beschikbare personeelsomvang, niet meer en niet minder), dan bereikt ook deze kostencategorie zijn minimale waarde, te weten nul. Zijn dat werkelijk optima? Neen immers! Het zijn slechts partiële optima, relevant wanneer we een enkele kostencategorie bezien maar bepaald niet wanneer het gaat om het totaal. Want neem eens aan, dat de fabrieksdirecteur de voorraad op de afzet zou afstemmen volgens $k_8 + k_9v$ en bovendien zowel overwerk als leegloop systematisch zou vermijden. Het resultaat is als volgt: iedere verandering van de afzet leidt dan noodzakelijkerwijs,

1. Gedurende de periode, waarvoor Holt c.s. hun studie uitvoerden, lag het maximale afzetniveau ruwweg 50% boven het gemiddelde en het minimale 50% lager.

gegeven de aan te passen voorraad, tot een bepaalde verandering van de productie; de productie is gekoppeld aan de personeelsbezetting volgens de formule $p = k_5q$, want overwerk en leegloop worden immers vermeden; dus leidt de verandering van de productie noodzakelijkerwijs tot een bepaalde verandering van de personeelsbezetting; en dat kost geld, hetzij omdat nieuw personeel moet worden aangetrokken hetzij omdat ontslagen moeten vallen. Kort gezegd, als de fabrieksdirecteur zijn aandacht zou beperken tot slechts twee van de kostencategorieën, nl. door voorraad- en omschakelingskosten enerzijds en overwerk- en leegloopkosten anderzijds beide tot een minimum te reduceren, dan bekoopt hij deze eenzijdigheid door de hogere kosten van de grote verandering van de personeelsomvang. En evenzo, wanneer de personeelsomvang constant wordt gehouden en overwerk- en leegloopkosten minimaal worden gehouden, dan is het gevolg een wild fluctueren van de voorraad met de daaraan verbonden kosten. De beste procedure is uiteraard die, welke het *totaal* van *alle* kosten minimeert. Hierop zullen wij ons nu richten. Daarbij treedt een aantal problemen op, die het eenvoudigst puntsgewijs kunnen worden behandeld.

(1) Het minimeren van de totale kosten vindt niet 'onvoorwaardelijk' plaats, omdat we rekening dienen te houden met een restrictie. Wanneer de productie in een bepaalde maand 100 eenheden bedraagt en er worden 80 voor verkoop afgeleverd, dan stijgt de voorraad met 20 eenheden. Algemener: de verandering van de voorraad is gelijk aan het exces van productie boven afzet. In formulevorm:

$$y - y_{-1} = p - v,$$

waarbij y_{-1} de voorraad van een maand tevoren voorstelt. Uiteraard kan de voorraadverandering $y - y_{-1}$ ook negatief zijn, dus op een verlaging van de voorraad neerkomen, nl. wanneer de productie beneden de afzet ligt.

(2) Aldus is ons probleem als volgt omgevormd: minimeer de totale kosten met inachtneming van de zojuist besproken restrictie inzake voorraadverandering. Maar daarmee zijn we er nog niet, want de vraag rijst: over welke termijn dienen die kosten te worden geminineerd? We hebben ons tot nu toe beperkt tot één enkele maand, maar het is duidelijk dat de fa-

brieksdirecteur bijzonder kortzichtig zou handelen door zich toe te leggen op de minimering van de kosten van die ene maand. Het zou immers neerkomen op een 'après nous le déluge'. Hij kan een horizon van bijv. 6 of 12 maanden aanhouden en zich ten doel stellen het totaal generaal van alle kosten over die periode te minimeren. Hij kan – als limietgeval – een oneindig lange horizon aanhouden; dat klinkt wellicht wat abstract, maar het is een redelijke procedure wanneer de fabrieksdirecteur de continuïteit van zijn bedrijf gedurende een onbepaalde periode op het oog heeft. Hoe dan ook, het is duidelijk dat de keuze van een horizon van meer dan een maand tweeërlei gevolg heeft. Enerzijds hebben we nu met een 'dubbele sommatie' te doen: we tellen niet alleen de vier kostencategorieën per maand op, maar we dienen bovendien de totale kosten per maand op te tellen om te komen tot het totaal generaal van alle kosten over de gehele periode, en onze doelstelling wordt het minimeren van dit totaal generaal. Anderzijds wordt het noodzakelijk onze variabelen te 'dateren'. We hebben nl. te doen met de productie in de maand die onmiddellijk voor ons ligt, met de productie 1 maand daarna, 2 maanden daarna, enz. We zullen die successieve productieniveaux met indices aangeven:

$p_0, p_1, p_2, p_3, \ldots,$

waarbij p_0 dus staat voor de productie nu, d.w.z. van de maand onmiddellijk voor ons. Analoog hebben we $q_0, q_1, \ldots$ voor de personeelsbezetting in de successieve maanden; $y_0, y_1, \ldots$ voor de voorraden en $v_0, v_1, \ldots$ voor de verkopen. We brengen in herinnering, dat q_{-1} en y_{-1} stonden voor de beginwaarden van personeelsbezetting resp. voorraad. Deze indexnotatie is geheel in overeenstemming met de hier geïntroduceerde. Immers, q_0 en y_0 staan voor personeelsbezetting resp. voorraad aan het einde van de maand die onmiddellijk voor ons ligt; gaan we een maand terug (zodat de indices 0 vervangen worden door —1), dan hebben we de personeelsbezetting en de voorraad aan het begin van de periode die we hier beschouwen.

(3) Onze probleemstelling is dus wederom gewijzigd. Hij luidt nu: minimeer het totaal van alle kosten over de gehele horizon, zulks met inachtneming van de restricties inzake voorraadverandering. 'Restricties' is opzettelijk in het meervoud gebruikt. We hebben nl. een restrictie voor elke maand: de

voorraadverandering in de eerste maand is gelijk aan het exces van productie boven afzet van de eerste maand, de voorraadverandering in de tweede maand is gelijk aan het analoge exces van de tweede maand, enz.

Maar we zijn er nog niet, want er is het probleem van onzekerheid. De voorraad- en omschakelingskosten zijn afhankelijk van de afzet en daarmee hangen de totale kosten af van de afzet. Wanneer dit de afzet van een zekere maand in het verleden zou zijn, dan was er geen probleem; men kan immers redelijkerwijs aannemen, dat de afzet van het verleden bekend is. Maar het gaat juist om de afzet van alle maanden van de horizon, dus om v_0, v_1, v_2, . . .; en dat is toekomst, dus die afzet is onbekend. Dit betekent, dat wanneer onze fabrieksdirecteur tot een bepaalde gedragslijn zou besluiten, hij niet in staat is van tevoren het daaraan verbonden kostenbedrag te bepalen. Hij is óók niet in staat om van tevoren die gedragslijn uit te stippelen die tot het kleinste totale kostenbedrag leidt. Welke gedragslijn deze optimale eigenschap heeft hangt af van het verloop van de afzet, en dat verloop is van tevoren niet bekend. Kennelijk maakt deze situatie het criterium van zo gering mogelijke totale kosten tot een weinig hanteerbare zaak. Wij zullen daarom dit criterium wijzigen: in hetgeen volgt zullen wij ons baseren op het minimeren van de *mathematische verwachting* van de totale kosten. Wil dit mogelijk zijn, dan moet de fabrieksdirecteur in staat zijn waarschijnlijkheidsuitspraken omtrent de toekomstige afzet te doen. Deze kunnen bijv. van het volgende type zijn:

> De afzet in de maand onmiddellijk vóór ons (v_0) is normaal verdeeld met een verwachting, die gelijk is aan 95% van het meest recente gerealiseerde niveau, en een standaarddeviatie gelijk aan 4% van dat niveau.
> De afzet in de maand daarna (v_1) is normaal verdeeld met een verwachting gelijk aan 96% van datzelfde niveau en een standaarddeviatie gelijk aan 6% van dat niveau.
> Enzovoorts.

In dit voorbeeld hebben we de standaarddeviatie laten oplopen (van 4 naar 6%). Dit ligt voor de hand, omdat de standaarddeviatie een maatstaf is voor de onzekerheid omtrent de toekomstige afzet en omdat we i.h.a. mogen verwachten, dat die

onzekerheid toeneemt naarmate de afzet op een latere toekomst betrekking heeft.

De zaak is nu rond wat de formulering van het criterium betreft. Het gaat om de totale kosten verbonden met productie, voorraad en personeelsbezetting, en wel de totale kosten over een zekere horizon. Het criterium is het minimeren van de mathematische verwachting van de totale kosten met inachtneming van de restricties inzake voorraadverandering. Maar hoe? Om de oplossingsprocedure goed te doorzien is het noodzakelijk onze variabelen in twee groepen te onderscheiden. We hebben enerzijds de variabelen, die de fabrieksdirecteur zelfstandig, autonoom beheerst. Dit geldt voor het productieniveau en voor de personeelsbezetting. Onze fabrieksdirecteur is immers in staat – althans, dat nemen we aan – de hoogte van de productie maand na maand vast te stellen en evenzo is hij in staat personeel aan te stellen en te ontslaan op door hem gewenste tijdstippen. Anderzijds hebben we de variabelen, die niet of onvolledig door de fabrieksdirecteur worden beheerst. Dit is uiteraard van toepassing op de afzet, waarvan de hoogte bepaald wordt door de afnemers. Het is ook van toepassing op de voorraad. Immers, hoe groot de voorraad is aan het eind van zekere maand wordt bepaald (i) door de voorraad aan het begin van de maand, (ii) door de productie gedurende die maand en (iii) door de afzet van die maand. Maar we concludeerden zojuist, dat die afzet buiten de macht van de fabrieksdirecteur ligt; en dus beheerst hij het verloop van de voorraad onvolledig.

Waarom is dit onderscheid zo belangrijk? Omdat onze fabrieksdirecteur slechts door aanpassing van de variabelen die hij beheerst aan zijn criterium kan voldoen! Aan de variabelen die hij niet beheerst kan hij niet veel meer doen dan hopen dat zij zich ordentelijk gedragen. Kort gezegd, hij beschikt over de productievariabelen in successieve maanden:

$p_0, p_1, p_2, p_3, \ldots$

en ook over de personeelsomvang:

$q_0, q_1, q_2, q_3, \ldots;$

en wat hij dient te doen is het aanpassen van al die p's en q's zodanig, dat de verwachte totale kosten geminimeerd worden met inachtneming van de restricties inzake voorraadverandering. Hoe? Door het vastprikken van de p's en q's op die nume-

rieke waarden waarvoor (met inachtneming van de bekende restricties) de verwachte totale kosten geminimeerd worden? Neen! Want wij weten van Hoofdstuk 7, dat de beste procedure een strategie is, een beslissingsregel, die specificeert welke besluiten successievelijk genomen dienen te worden afhankelijk van de informatie, die beschikbaar zal zijn op het moment waarop de beslissing in feite genomen moet worden. In het onderhavige geval bestaat die informatie uit het verloop van de afzet. Nu, op dit moment, kent de fabrieksdirecteur de afzet van het verleden; bovendien heeft hij zekere ideeën over de afzet in de successieve toekomstige perioden, nl. in de vorm van waarschijnlijkheidsverdelingen. Een maand later kent de fabrieksdirecteur de afzet van de maand, die op dit moment onmiddellijk vóór hem ligt (v_0) en dus nu nog toekomst is; bovendien kan hij zekere aanwijzingen gekregen hebben (bij verf: bijv. het aanbesteden van nieuwe gebouwen), die zijn oorspronkelijke ideeën omtrent de afzet van latere maanden doen veranderen. Kortom, gedurende die maand zal hij informatie ontvangen, waarvan hij gebruik dient te maken voor zijn beslissing inzake productie en personeelsbezetting aan het begin van de volgende maand (p_1 en q_1). En dit wordt keer op keer herhaald: een maand later kent hij de afzet van die maand (v_1), de waarschijnlijkheidsverdelingen van de afzet in latere maanden zijn wederom veranderd, deze nieuwe informatie wordt gebruikt voor p_2 en q_2, enz.

5. LINEAIRE BESLISSINGSREGELS

Het gaat dus om de beslissingsregel, die de verwachting van de totale kosten minimeert met inachtneming van de restricties inzake voorraadverandering. Nu zijn die *restricties* alle *lineair* (in voorraad, productie en afzet). Verder is de *kostenfunctie kwadratisch.* Hij is immers de som van een viertal kostencategorieën (bovendien nog gesommeerd over de tijd), die alle van de eerste of de tweede graad in de diverse variabelen zijn.

Wanneer men te doen heeft met een kwadratische doelstellingsfunctie, waarvan de verwachting over de tijd geminimeerd of gemaximeerd dient te worden met inachtneming van lineaire restricties, dan treden een aantal vereenvoudigingen op, die de hanteerbaarheid van de oplossing bijzonder ten goede komen.

Het gaat hierbij vooral om de beslissing die nu, op dit moment, moet worden genomen; dus voor het verfprobleem de hoogte van de productie en de omvang van het personeel in de maand die onmiddellijk vóór ons ligt (p_0 en q_0). Nog specifieker: het gaat om de p_0 en de q_0 van de beste strategie, dus van de strategie die de verwachte totale kosten minimeert met inachtneming van de lineaire restricties inzake voorraadverandering. Het blijkt nu (het bewijs wordt hier niet gegeven), dat deze p_0 en q_0 op een eenvoudige manier kunnen worden berekend, nl. door simpelweg alle onzekerheid weg te denken. Preciezer, wanneer de fabrieksdirecteur de enige bron van onzekerheid in zijn probleem, dus die van zijn afzet in toekomstige perioden, elimineert door die afzet door de mathematische verwachting ervan te vervangen; en wanneer hij vervolgens de totale kosten minimeert (met inachtneming van de restricties inzake voorraadverandering) en daarbij doet alsof de toekomstige afzet inderdaad samenvalt met die mathematische verwachting; dan komt hij aldus doende automatisch op precies dezelfde productie- en personeelsomvangbeslissing voor de maand die onmiddellijk vóór hem ligt als de p_0 en q_0 van de beste strategie. Als voorbeeld nemen we de waarschijnlijkheidsuitspraken omtrent de toekomstige afzet van blz. 288. Daar werd gesteld, dat v_0 (de afzet in de maand onmiddellijk vóór ons) normaal verdeeld is met een verwachting gelijk aan 95% van het meest recente niveau en een standaarddeviatie van 4% van dat niveau; voorts dat v_1 ook normaal verdeeld is, maar met een verwachting van 96% en een standaarddeviatie van 6% van dat niveau. Welnu, het zojuist genoemde resultaat komt dan hierop neer, dat de p_0 en de q_0 van de beste strategie gevonden kunnen worden door te doen alsof de toekomstige afzet v_0 samenvalt met zijn mathematische verwachting (dus het 95% niveau), alsof v_1 samenvalt met zijn verwachting (96%), en analoog voor v_2, v_3, ...

Dit resultaat staat bekend als 'zekerheidsequivalentie'. D.w.z., door te doen alsof de toekomstige afzet met zekerheid gelijk is aan zijn mathematische verwachting verkrijgen we de beslissingen p_0 en q_0 van de beste strategie. Uiteraard 'doen alsof'; de toekomstige afzet zal, na een of meer maanden later te zijn gerealiseerd, vrijwel zeker niet exact blijken samen te vallen met de verwachting. We hebben immers met spreiding te doen; er is een standaarddeviatie van 4% voor v_0 en van 6% voor v_1.

Maar daar gaat het niet om. Het moge volledig correct zijn, dat de standaarddeviatie de neiging heeft op te lopen naarmate we verder in de toekomst gelegen perioden bezien (zoals op blz. 288 werd opgemerkt); dit is echter volmaakt overbodige informatie wanneer het om de p_0 en de q_0 van de beste strategie gaat, omdat die immers uitsluitend door de verwachting van de afzet bepaald worden. We kunnen de standaarddeviaties van 4% en 6% vervangen door bijv. 2% resp. 8% en dat maakt niets uit voor die p_0 en q_0, zolang de verwachte afzet in de successieve maanden maar blijft wat hij is.

Maar hoe worden die p_0 en q_0 nu in concreto bepaald? Welnu, het spreekt vanzelf dat die afhankelijk zullen zijn van de verwachte afzet in de successieve maanden, weer te geven door Ev_0, Ev_1, Ev_2, ... Aangetoond kan worden, dat die afhankelijkheid van het lineaire type is; vandaar dat men spreekt over *lineaire beslissingsregels*. Dit is naast zekerheidsequivalentie een tweede vereenvoudigende omstandigheid bij het opsporen van de p_0 en de q_0 van de strategie, die de verwachting van een kwadratische kostenfunctie over de tijd minimeert met inachtneming van lineaire restricties. Daarnaast zijn er nog de 'beginvoorwaarden'. De fabrieksdirecteur neemt immers zijn beslissing vanuit een bestaande situatie met een gegeven beginvoorraad (y_{-1}) en een gegeven personeelsomvang in het begin (q_{-1}). Het ligt in de rede, dat de beslissingsregel voor p_0 ons zal vertellen dat dit productieniveau lager is naarmate de fabriek al beschikt over een grotere beginvoorraad. Dat is inderdaad het geval, zoals we nu onmiddellijk zullen zien.

Holt c.s. hebben de beslissingsregels voor p_0 en q_0 numeriek gespecificeerd voor de hier beschouwde verffabriek. Daartoe hebben ze eerst de coëfficiënten (de k's) van de vier kostencategorieën bepaald met behulp van interne gegevens van het bedrijf; verder hebben ze zich bepaald tot het geval van een oneindige horizon, zodat de fabrieksdirecteur dus verondersteld wordt een onbepaald lange tijd vooruit te kijken. Het resultaat voor de beslissingsregel van het productieniveau is als volgt:

$$p_0 = \left\{ \begin{array}{l} +0{,}464Ev_0 \\ +0{,}236Ev_1 \\ +0{,}112Ev_2 \\ +0{,}047Ev_3 \\ +0{,}014Ev_4 \end{array} \right\} - 0{,}464y_{-1} + 1{,}007q_{-1} + 153{,}1.$$

Deze beslissingsregel beschrijft het productieniveau volgens de

beste strategie voor de eerstkomende maand als de som van een drietal uitdrukkingen. Van achter naar voren bezien hebben we in de eerste plaats een constante term (153,1). In de tweede plaats zijn er de beginvoorwaarden. De regel specificeert, dat als de beginvoorraad y_{-1} één eenheid groter zou zijn, het toe te passen productieniveau 0,464 eenheden lager wordt; hetgeen wat de richting van het effect betreft in overeenstemming is met het vermoeden, dat aan het einde van de vorige alinea werd uitgesproken. Voorts zien we, dat bij een grotere personeelsbezetting aan het begin van de maand het toe te passen productieniveau wordt verhoogd; dit volgt uit de positieve coëfficiënt (1,007) van q_{-1}. De oorzaak van dit effect is gelegen bij de overwerk- en leegloopkosten en bij de kosten van personeelsmutaties. De overwerk- en leegloopkosten zijn nl. minimaal wanneer de fabriek precies op capaciteit werkt, hetgeen een evenredigheid tussen p_0 en q_0 impliceert ($p_0 = k_5 q_0$). Wanneer er een grotere personeelsbezetting aan het begin van de maand (q_{-1}) is en wanneer die bezetting gedurende die maand niet verandert, dus $q_0 = q_{-1}$, dan is ook q_0 groter, zodat de overwerk- en leegloopkosten een stimulans voor een verhoogde p_0 betekenen, juist vanwege die evenredigheid. Vandaar de positieve coëfficiënt van q_{-1}. Nu is het bepaald niet noodzakelijk, dat de personeelsbezetting gedurende de maand onveranderd blijft, zoals we hier even gemakshalve hebben aangenomen. Maar als die wél verandert zijn daar kosten aan verbonden. Vandaar dat de positieve coëfficiënt van q_{-1} uiteindelijk aan een tweetal kostencategorieën moet worden toegeschreven.

In de derde plaats tenslotte hebben we de termen in de verwachte afzet van de successieve maanden. Wanneer de verwachte afzet van de eerstkomende maand (Ev_0) met één eenheid stijgt, wordt het toe te passen productieniveau verhoogd met bijna een halve eenheid (preciezer: met 0,464 eenheden). Kennelijk wordt de extra eenheid afzet niet onmiddellijk geproduceerd, althans niet meer dan halverwege. Zou dit wel gedaan worden, dan kost dit overwerk of er moeten employé's worden aangetrokken, hetgeen ook weer geld kost; blijkbaar is het beter die extra eenheid gedeeltelijk ten koste van de voorraad te laten gaan. De invloed van de verwachte afzet één maand later is ongeveer de helft van die van de eerstkomende maand; in de daarop volgende maand wordt die invloed weer

ongeveer gehalveerd, daarna loopt het sneller af. In beginsel loopt de reeks ook na Ev_4 door (en wel tot in het oneindige toe); we hebben ons hier beperkt tot termen met een coëfficiënt boven 0,01.

Ook voor de personeelsbezetting is er een beslissingsregel. Deze luidt als volgt:

$$q_0 = \left\{ \begin{array}{l} +0{,}0100Ev_0 \\ +0{,}0087Ev_1 \\ +0{,}0070Ev_2 \\ +0{,}0054Ev_3 \\ +0{,}0041Ev_4 \\ +0{,}0030Ev_5 \\ +0{,}0022Ev_6 \\ +0{,}0016Ev_7 \\ +0{,}0011Ev_8 \end{array} \right\} - 0{,}0100y_{-1} + 0{,}742q_{-1} + 2{,}00.$$

Deze beslissingsregel bevat dezelfde termen als die voor het productieniveau, echter met andere coëfficiënten. Ook hier hebben we een constante term (2,00) en twee termen voor de beginvoorwaarden. De coëfficiënt van de personeelsbezetting aan het begin van de maand (0,742) is zeer aanzienlijk. Dit moet in de eerste plaats toegeschreven worden aan de kosten verbonden aan personeelsmutaties en het resultaat is, ruwweg gesproken, dat bijna 75% van die personeelsbezetting aan het einde van de maand vastligt door de omvang van de bezetting aan het begin van de maand. De coëfficiënt van de beginvoorraad (y_{-1}) is negatief; allicht, want een grotere beginvoorraad maakt het minder noodzakelijk om mensen aan het werk te zetten om nog meer te produceren. Tenslotte: de verwachte afzet. Evenals bij de beslissingsregel voor de productie heeft de verwachte afzet een positieve invloed op de te realiseren personeelsomvang en daalt die invloed naarmate we verder in de toekomst kijken. Maar de daling is nu veel langzamer. Bij de productie was het een kwestie van successieve halvering, hier is de coëfficiënt van Ev_1 bijna 90% van die van Ev_0, die van Ev_2 is ongeveer 80% van die van Ev_1, enz.[1] Waarom dit verschil? Wel, stel dat de verwachte afzet in de successieve maanden nogal

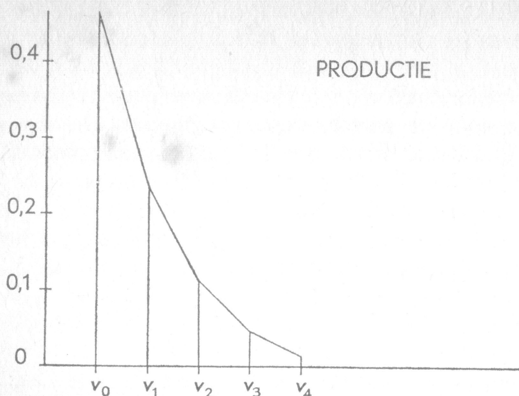
1. Ook voor de beslissingsregel van de productieomvang geldt, dat er in beginsel oneindig veel termen van afzetverwachtingen zijn. Wij hebben de reeks afgekapt op het punt, waar de coëfficiënten kleiner dan 0,001 worden.

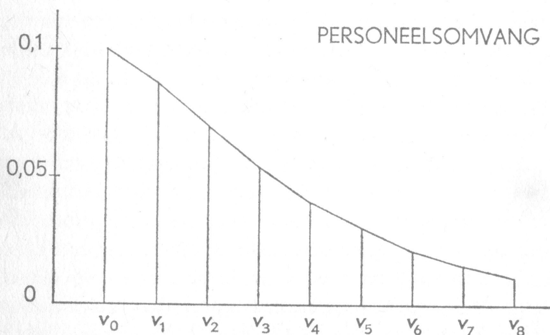
wild fluctueert. Dan is het nog altijd voordeliger (gegeven de kostenstructuur van de fabriek) de productie op korte termijn aan die fluctuaties aan te passen dan de personeelsbezetting. Uiteraard past de productie zich niet onmiddellijk en volledig aan (zie de coëfficiënt 0,464 van de desbetreffende beslissingsregel), maar voor de personeelsbezetting wordt dit nog bezwaarlijker in verband met de specifieke kosten aan personeelsmutaties verbonden. Vandaar dat de beslissingsregel voor q_0 de tendentie heeft wilde fluctuaties van de verwachte afzet 'uit te smeren': de successieve verwachtingen Ev_0, Ev_1, ... hebben langzaam aflopende coëfficiënten, zodat een paar uitschieters in de rij betrekkelijk weinig invloed hebben. Althans veel minder invloed dan bijv. in het geval van een uitschietende Ev_0 bij de beslissingsregel voor de productie.

Een duidelijk beeld van het verschil tussen de twee reactiepatronen – voor productie resp. personeelsomvang, beide t.o.v. de successieve afzetverwachtingen – verkrijgt men met behulp van een grafiek. Langs de horizontale as zetten we op gelijke afstanden de verwachtingen van v_0, v_1, ... af en langs de verticale assen (twee stuks: één voor productie, één voor de personeelsomvang) de numerieke waarden van de bijbehorende coëfficiënten. Die coëfficiënten worden dan gerepresenteerd door verticale staven, elk ter lengte van de respectieve numerieke waarde. Het resultaat treft men aan op de volgende bladzijde. Om het verloop van de successieve coëfficiënten nog iets beter te laten uitkomen hebben we de eindpunten van de staven successievelijk door rechte lijnstukken verbonden. Duidelijk is, dat de daling bij de productie veel sterker is dan bij de personeelsomvang.

6. CONCRETE GEVALLEN

Het is instructief om te zien hoe deze beslissingsregels in een eenvoudig geval in concreto werken. Daartoe beschouwen we de volgende situatie: de afzet is constant op een vast niveau, echter met uitzondering van een uitschieter van zegge 100 eenheden in één bepaalde maand. Teneinde de zaak niet te gecompliceerd te maken zullen wij in eerste aanleg veronderstellen, dat dit verloop van de afzet van tevoren exact bekend is. Dit zal bijv. het geval zijn, wanneer de afnemers vaste contracten



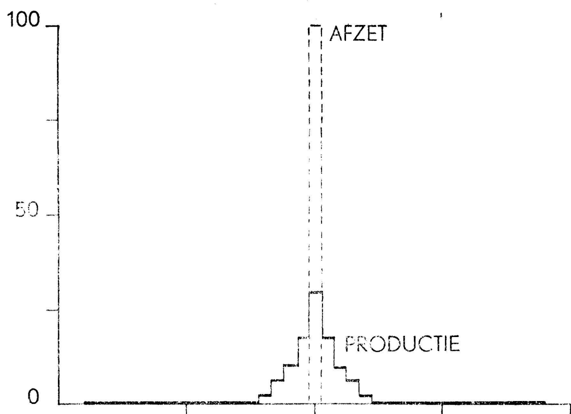


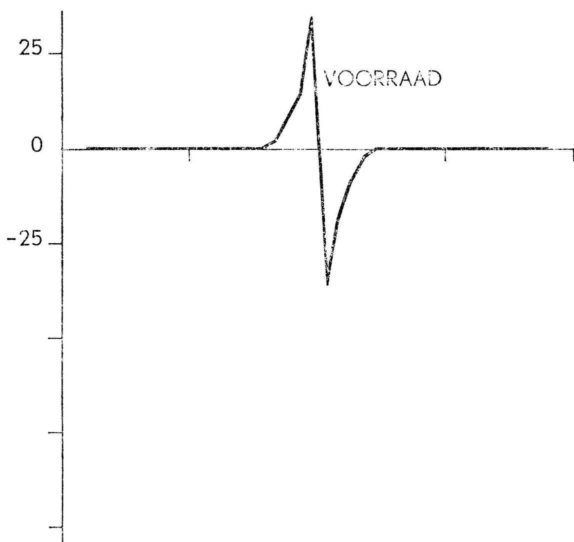
tot constante aflevering hebben en wanneer die extra order van 100 eenheden van tevoren is aangekondigd.

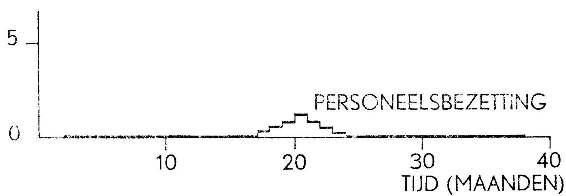
Het probleem is: hoe dient de fabrieksdirecteur wat productie en personeelsomvang betreft op dit afzetverloop te reageren? Het antwoord is simpel: dat kan hij aflezen uit zijn beslissingsregels! De productieregel vertelt hem dat hij vier maanden vooruit $100 \times 0{,}014 = 1{,}4$ extra eenheden dient te produceren; immers, het gaat dan om een afzet van 100 eenheden 4 maanden later en de coëfficiënt van Ev_4 is 0,014. Een maand later, dus drie maanden vooruit, zal hij $100 \times 0{,}047 = 4{,}7$ extra eenheden produceren, enz. Voor de personeelsomvang is de procedure analoog. Het gevolg is, dat vóór de 100 extra eenheden

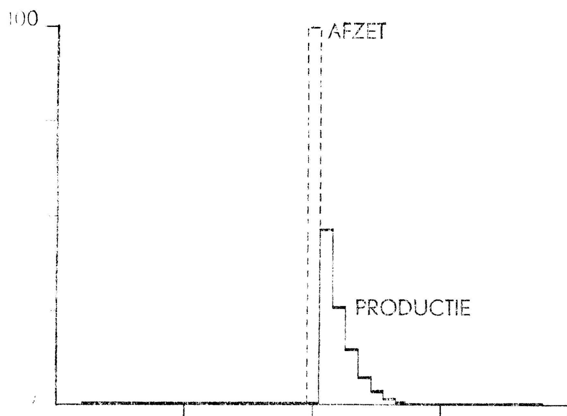
worden afgeleverd, er meer wordt geproduceerd dan verkocht; de voorraad stijgt dus. In de maand waarin die extra eenheden worden afgeleverd daalt de voorraad plotseling, en wel beneden het oorspronkelijke niveau. De productie daalt ook, maar geleidelijk; hij blijft nog enkele maanden boven zijn oorspronkelijke niveau, zodat de voorraad gelegenheid krijgt weer aan te groeien. Inmiddels is ook de personeelsbezetting van zijn tijdelijke top naar het oorspronkelijke niveau teruggekeerd. Het totale beeld in het verloop van de tijd vindt men op blz. 298 getekend. De horizontale assen van deze grafieken zijn aangebracht op het evenwichtsniveau van vóór de uitschietende afzet; de schaalverdelingen van de verticale assen meten dus de desbetreffende variabelen in afwijking van dit evenwichtsniveau. Duidelijk blijkt uit de drie grafieken, dat deze afwijkingen een a.h.w. symmetrisch verloop hebben t.o.v. de maand van de uitschieter.

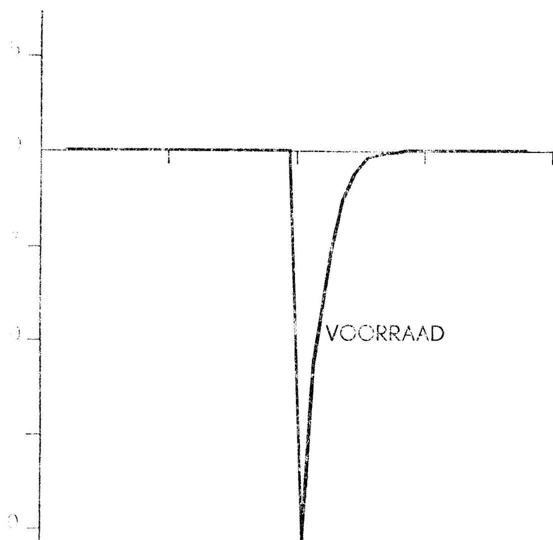
Als tweede voorbeeld nemen we dezelfde situatie, echter met dit verschil dat de fabrieksdirecteur tevoren geen weet heeft van de extra order van 100 eenheden. Dit betekent, dat zijn verwachte afzet nu constant is, maand na maand, op het evenwichtsniveau. De uitschietende afzet komt dan als een volledige verrassing. Hoe zal het verloop nu zijn? Dit is getekend op blz. 299 naast het zo juist behandelde geval en de verklaring is als volgt. Plotseling komt die extra order binnen. Daar is tevoren niet op gerekend; in tegenstelling tot het eerste geval liggen alle variabelen dus op de horizontale as (het evenwichtsniveau) vóór de maand van de uitschieter. Dan wordt opeens de voorraad met 100 eenheden gereduceerd vanwege de extra order. Aan het eind van de maand is dit de beginvoorraad y_{-1}. Volgens de beslissingsregel voor de productie moet het productieniveau van die maand met $100 \times 0{,}464 = 46{,}4$ eenheden worden verhoogd; immers, y_{-1} is verhoogd met —100 en zijn coëfficiënt is —0,464. Dit is het enige effect wat p_0 betreft, want de afzetverwachtingen voor latere maanden zijn onveranderd en ook de personeelsbezetting aan het begin van de maand is als vanouds. De productieverhoging van 46,4 eenheden heft de reductie van de voorraad bijna halverwege op; tegelijkertijd wordt ook de personeelsbezetting opgevoerd. Wanneer we de twee tekeningen vergelijken, valt vooral op dat de voorraadafwijking veel groter is geworden – en dat kost geld. Het zal intuïtief duidelijk zijn dat de totale kosten in het

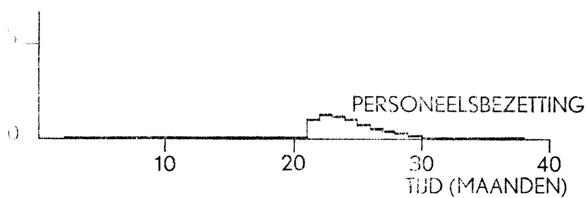

100
50
0
AFZET
PRODUCTIE


25
0
-25
VOORRAAD


5
0
PERSONEELSBEZETTING
10
20
30
40
TIJD (MAANDEN)


100
AFZET
PRODUCTIE


VOORRAAD


PERSONEELSBEZETTING
10
20
30
40
TIJD (MAANDEN)

onderhavige geval hoger moeten zijn dan in het voorgaande.

Nu het derde voorbeeld – wellicht het meest voor de hand liggende. Stel dat de verffabriek waarom het hier gaat die beslissingsregels werkelijk in practijk had gebracht; zouden de aandeelhouders er beter op zijn geworden? Dat is een toets waarvan de practische waarde toch niet in twijfel kan worden getrokken! Ook dit is door Holt c.s. onderzocht. Zij hebben nl. voor een periode in het verleden van drie jaren de door de fabriek in feite gemaakte kosten vergeleken met die, welke gemaakt zouden zijn als de beslissingsregels waren toegepast. Een tweetal kleine korrels zout is bij een dergelijk onderzoek overigens wel op zijn plaats. Ten eerste doen zich bij kostenbepaling nogal eens problemen voor, die de allure van Gordiaanse knopen hebben; wil men met een eenduidig numeriek resultaat voor de dag komen, dan is de hantering van het bekende zwaard vaak onvermijdelijk, met alle bezwaren van dien. Ten tweede plegen fabrieksdirecteuren niet steeds maand na maand hun afzetvoorspellingen op schrift te stellen. Dit laatste probleem werd aangepakt door met twee alternatieve veronderstellingen te werken: een optimistische, waarbij ervan wordt uitgegaan dat het verloop van de afzet tevoren volmaakt was voorspeld, en een pessimistische, die neerkomt op de gedachte dat de fabrieksdirecteur tot niet meer in staat is dan een zuiver-mechanische voorspellingsmethode.

Zo komen we tot drie kostenopstellingen: de feitelijke, die het gedrag van de fabrieksdirecteur beschrijft zoals dit werkelijk heeft plaats gevonden; en twee volgens de beslissingsregels, één optimistisch en één pessimistisch. Drukken we de kosten uit in duizenden dollars, dan ontstaat het volgende beeld:

		Beslissingsregel	
	Feitelijk	Pessimistisch	Optimistisch
Normale loonkosten	1940	1834	1888
Aantrekken en afstoten	22	25	20
Overwerk en leegloop	196	296	167
Voorraadkosten	361	451	454
Achterstallige leveringen	1566	616	400
Totaal	4085	3222	2929

De totale kosten zijn het geringst, zoals men verwachten zou, bij toepassing van de beslissingsregels onder de veronderstelling dat volmaakt wordt voorspeld. De mechanische voorspelling reduceert het voordeel van de beslissingsregels uiteraard; de gemaakte kosten zijn dan ongeveer 10% hoger. De in feite gemaakte kosten zijn nog weer ruim 25% hoger dan dit bedrag, hetgeen suggereert dat toepassing van de beslissingsregels als zodanig belangrijker is dan volmaakt voorspellen. Beziet men de opstelling meer in detail, dan valt op dàt de verschillen vooral liggen bij de kosten verbonden aan de achterstallige leveringen (die daarom afzonderlijk naast de eigenlijke voorraadkosten zijn opgevoerd). De beslissingsregels geven te zien, dat het aanbeveling had verdiend grotere voorraden aan te houden, zodat de achterstallige leveringen daardoor geringer geweest zouden zijn. Gedeeltelijk ligt dit aan de bijzondere aard van de onderzochte periode: deze omvat nl. het jaar van de Koreaanse oorlog. Een herhaling van het onderzoek voor een latere periode, die het Korea-jaar niet omvat, laat zien dat de kosten van de achterstallige leveringen dan niet meer zo excessief zijn; de beslissingsregels onder de pessimistische voorspellingsveronderstelling zijn dan niet meer zo veel voordeliger dan wat er in feite gebeurd is, al blijft er dan nog een verschil in totale kosten over van een kleine 10%.

7. VOORSPELLINGEN VAN TOEKOMSTIGE BESLISSINGEN

In de zojuist behandelde toepassingen is de beslissingsregel maand na maand toegepast, zodat de betekenis van p_0 en q_0 iedere maand verschoof: begin januari zijn dit de beslissingen inzake productie en personeelsomvang voor januari, een maand daarop voor februari, enz. Daarmee is in feite de vraag beantwoord, hoe men in concreto zich conform de beste strategie dient te gedragen. Desondanks is het van belang hieraan toe te voegen, dat de berekeningsprocedure in feite nog andere interessante resultaten oplevert; men kan nl. een voorspelling maken van het toekomstige gedrag volgens de beste strategie in de latere maanden. Teneinde dit aanschouwelijk te maken roepen we in herinnering, dat de beslissing voor nu, dus voor de maand die onmiddellijk vóór ons ligt, als volgt werd afgeleid: we minimeren de totale kosten (met inachtneming van de restric-

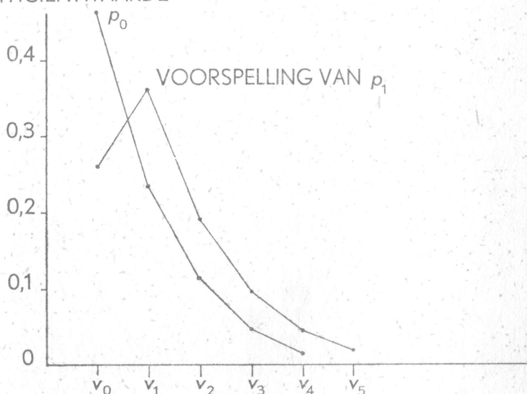
ties inzake voorraadverandering), na de onzekere toekomstige afzet vervangen te hebben door de mathematische verwachting. De uitkomst voor de eerstkomende maand is dan – volgens de stelling van zekerheidsequivalentie – gelijk aan de p_0 en q_0 van de beste strategie. Maar dezelfde minimeringsprocedure levert óók een uitkomst voor de eerste maand daarna. Zou dit de p_1 en q_1 van de beste strategie zijn? Neen, want de procedure houdt geen rekening (en kan ook geen rekening houden) met de informatie, die in de tussentijd – dus gedurende de eerstkomende maand – ter beschikking komt. Er is dus een verschil tussen enerzijds de beslissingen volgens de beste strategie één maand na dato en anderzijds de corresponderende waarden die de berekeningsprocedure in het heden daarvoor oplevert. Toch hebben die laatste waarden wel degelijk betekenis. In de eerste plaats zijn ze onmiddellijk beschikbaar, terwijl de p_1 en q_1 van de beste strategie pas een maand later berekend kunnen worden. In de tweede plaats kan, naar valt aan te tonen, een waarschijnlijkheidsuitspraak over het verschil worden gedaan. Het verschil nl. tussen enerzijds de p_1 van de beste strategie en anderzijds de waarde die daarvoor in het heden wordt berekend is soms positief en soms negatief, maar de mathematische verwachting van het verschil is nul; en analoog voor de corresponderende beslissing inzake de personeelsbezetting, q_1. We kunnen dus de in het heden berekende waarden als *voorspellingen* beschouwen van de p_1 en q_1 van de beste strategie; en wel als *zuivere* voorspellingen, hetgeen de technische uitdrukking is voor voorspellingen die de eigenschap hebben dat de voorspellingsfout (d.i. bovengenoemd verschil) verwachting nul heeft.

In het onderhavige geval vinden we de voorspelling van p_1 door toepassing van de volgende formule:

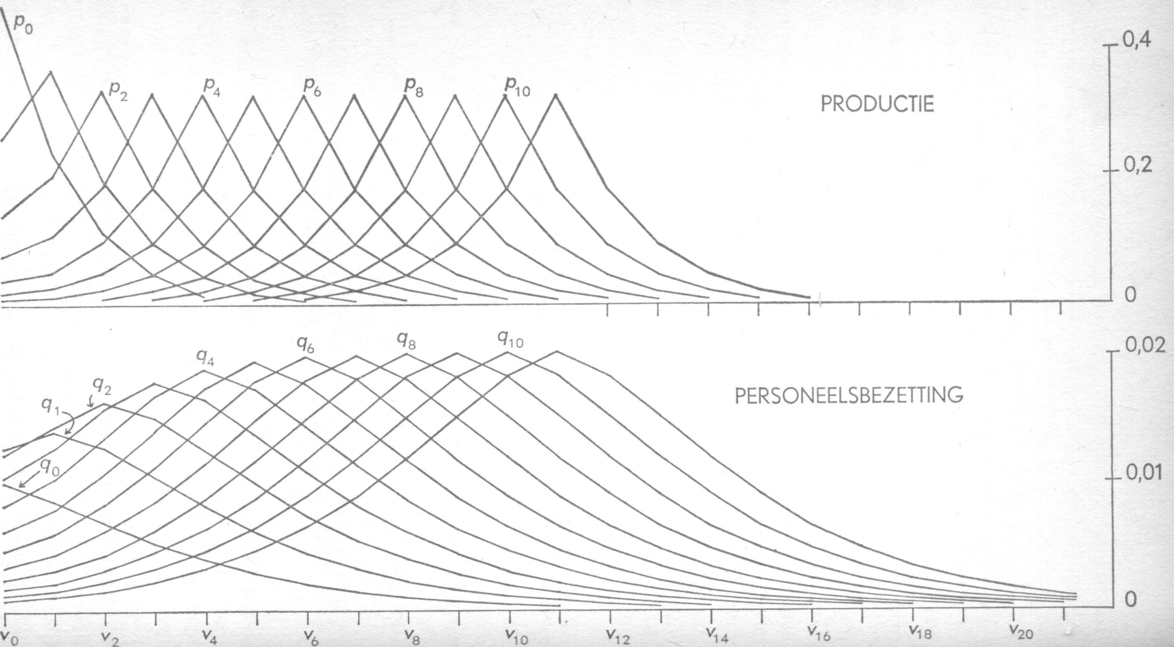
$$\text{Voorspelling van } p_1 = \left\{\begin{array}{l} +0{,}259Ev_0 \\ +0{,}363Ev_1 \\ +0{,}191Ev_2 \\ +0{,}096Ev_3 \\ +0{,}044Ev_4 \\ +0{,}018Ev_5 \end{array}\right\} - 0{,}259y_{-1} + 0{,}280q_{-1} + 84{,}1;$$

dus geheel analoog aan de beslissingsregel voor p_0, echter met andere coëfficiënten. In de eerste plaats liggen de coëfficiënten van de beginvoorwaarden (—0,259 en 0,280) dichter bij nul dan het geval was bij de vergelijking voor p_0. Dat ligt voor de hand,

want het niveau van voorraad en personeelsbezetting aan het begin van de eerste periode zijn voor de handelingen in de tweede periode van geringer belang dan voor die in de eerste periode. In de tweede plaats hebben de coëfficiënten van de verwachte afzet een ander verloop dan die in de vergelijking voor p_0. De grootste coëfficiënt is die van de verwachting van v_1, dus van de afzet in *dezelfde* maand als de maand waarvoor de productie (p_1) voorspeld wordt; ook dit ligt voor de hand, want door bij de vaststelling van p_1 een groot gewicht toe te kennen aan de afzet van dezelfde maand reduceren we de voorraadfluctuaties. Beide coëfficiëntenreeksen tezamen, dus enerzijds die van de verwachte afzet in de beslissingsregel voor p_0 en anderzijds die van de verwachte afzet in de voorspellingsvergelijking van p_1, kunnen op overzichtelijke wijze weergegeven en vergeleken worden door ze af te zetten in een diagram van hetzelfde type als dat van blz. 296:



We zien duidelijk, dat de coëfficiëntenreeks van p_0 een top heeft bij v_0, en die van p_1 bij v_1, en dat rond de top een geleidelijke daling plaats heeft. Dit nu geldt nog algemener. We kunnen nl. ook voorspellingen maken voor de productie in de tweede maand (p_2), in de derde (p_3), enz. De bijbehorende coëfficiëntenreeksen van de verwachte afzet vindt men in de bovenste helft van de figuur op blz. 304 afgebeeld. Duidelijk blijkt, dat die reeksen geleidelijk naar rechts verschuiven: de top van de coëfficiëntenreeks van p_2 ligt bij v_2, die van p_3 bij



PRODUCTIE
p_0
p_2
p_4
p_6
p_8
p_{10}
0,4
0,2
0
PERSONEELSBEZETTING
q_0
q_1
q_2
q_4
q_6
q_8
q_{10}
0,02
0,01
0
v_0
v_2
v_4
v_6
v_8
v_{10}
v_{12}
v_{14}
v_{16}
v_{18}
v_{20}

v_3, enz. We hebben hetzelfde beeld voor de personeelsbezetting, zie de onderste helft van de figuur. Daar echter is het verloop veel vlakker, hetgeen in overeenstemming is met de bijzonder langzame daling van de coëfficiënten van de beslissingsregel voor q_0, zie blz. 294.

Aldus zien we, dat het de fabrieksdirecteur mogelijk is om niet alleen te berekenen wat hem onmiddellijk te doen staat (p_0 en q_0) maar ook te voorspellen wat hij *zal* gaan doen in de komende maanden (dus voorspellingen van $p_1, q_1, p_2, q_2, \ldots$ – uiteraard geen volmaakte voorspellingen!). De benodigde ingrediënten zijn de beginvoorwaarden (y_{-1}: de beginvoorraad, en q_{-1} : de personeelsbezetting aan het begin); verder de mathematische verwachting van de toekomstige afzet (Ev_0, Ev_1, ...); en tenslotte de coëfficiënten waarmee die verschillende factoren moeten worden vermenigvuldigd, dus in feite de beslissingsregels voor p_0 en q_0 alsmede de voorspellingsvergelijkingen voor $p_1, q_1, p_2, q_2, \ldots$. Het resultaat is dan, dat we bijv. op 1 januari de beslissingen p_0 en q_0 van januari kunnen berekenen; verder ook voorspellingen van de februari-beslissingen p_1 en q_1, van de maart-beslissingen p_2 en q_2, enz. Merk op dat die voorspellingen elke maand *herzien* worden. Immers, op 1 februari, dus een maand later, zijn p_0 en q_0 de beslissingen voor februari, die een maand eerder als p_1 resp. q_1 werden aangemerkt. Op dat tijdstip (1 februari) worden de werkelijke februari-beslissingen berekend alsmede herziene voorspellingen voor maart, april, ...; en zo gaat het door.

LITERATUUR

De vierkantswortelformule voor de optimale seriegrootte is in 1915 door F. Harris gevonden. Verscheidene uitbreidingen van deze formule voor meer ingewikkelde gevallen zijn later afgeleid. Hierover (en over talrijke andere zaken betreffende voorraadbeleid) kan men lezen in het recente boek van Hadley en Whitin [1]. Verscheidene Nederlanders zijn actief op dit terrein, o.a. J. Goudriaan, W. Monhemius (hoogleraren resp. in Pretoria en in Eindhoven), A. R. van den Burg, R. N. van Hees en H. W. van den Meerendonk.

De verffabriek en zijn kostenstructuur zijn besproken door Holt e.a. [2]. De eerste stap op het terrein van zekerheidsequivalentie is gezet door Theil [3]; dit betrof het zgn. statische geval, waarin slechts voor één periode wordt geminimeerd of gemaximeerd. Voor de dynamische generalisatie zie Simon [4] en Theil [5]. Een uitvoerige

behandeling vindt men in Theil [6], waarin o.a. verslag wordt gedaan van onderzoekingen van C. van de Panne en van P. J. M. van den Bogaard en A. P. Barten. De eerstgenoemde richtte zich i.h.b. op de geldelijke verliezen, die worden geleden wanneer de directeur van de verffabriek beslissingen neemt die afwijken van de in de tekst behandelde 'beste' p_0 en q_0; wanneer de afzetvoorspellingen onvolmaakt zijn; en wanneer de coëfficiënten van de kostenstructuur (de k's) onjuist gespecificeerd zijn. De beide laatstgenoemden voerden een uitvoerig macro-economisch onderzoek uit. Het betreft overheidsbeslissingen voor de Nederlandse volkshuishouding in de periode 1957-1959. Negen variabelen speelden een rol in elk van die drie jaren, t.w. vijf variabelen die door de overheid werden beheerst (loonvoet, overheidsuitgaven, indirecte belastingen en directe belastingen geheven resp. op looninkomens en op overige inkomens) en vier variabelen die niet door de overheid werden beheerst (werkgelegenheid, prijspeil van consumptiegoederen, het aandeel van de lonen in het nationale inkomen en het saldo op de lopende rekening van de betalingsbalans). Deze negen variabelen werden onderling verbonden geacht door de vergelijkingen van een econometrisch macro-model van het type dat in Hoofdstuk 4 is beschouwd. Dit model stelt dan de restricties voor met inachtneming waarvan een kwadratische maatschappelijke voorkeursfunctie wordt gemaximeerd. Drie van dergelijke voorkeursfuncties zijn beschouwd: één representerend de voorkeuren van de werkgeversleden van de Sociaal-Economische Raad, één voor de werknemersleden en één voor de kroonleden.

[1] Hadley, G., and T. M. Whitin, *Analysis of Inventory Systems*. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1963.

[2] Holt C. C., F. Modigliani, J. F. Muth, and H. A. Simon, *Planning Production, Inventories, and Work Force*. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1960.

[3] Theil, H., 'Econometric Models and Welfare Maximization'. *Weltwirtschaftliches Archiv*, Vol. 72 (1954), pp. 60-83.

[4] Simon, H. A., 'Dynamic Programming under Uncertainty with a Quadratic Criterion Function'. *Econometrica*, Vol. 24 (1956), pp. 74-81.

[5] Theil, H., 'Note on Certainty Equivalence in Dynamic Economic Planning'. *Econometrica*, Vol. 25 (1957), pp. 346-349.

[6] Theil, H., *Optimal Decision Rules for Government and Industry*. Noord-Hollandsche Uitgevers Maatschappij, Amsterdam. 1964.

12. DE STATISTISCHE SPECIFICATIE VAN ECONOMISCHE RELATIES

1. HET PROBLEEM; DE TWEE BRONNEN VAN INFORMATIE

Op verscheidene plaatsen van dit boek hadden we te maken met het probleem van de vaststelling van het numerieke verband tussen twee of meer economische variabelen. Soms gebeurt dat op een in beginsel heel eenvoudige manier, zoals de relatie tussen onderlinge leveringen en totale productie per sector in de aan- en afvoeranalyse; we werkten daar met een veronderstelde evenredigheid en konden dan volstaan met een tabel van een enkel jaar. Soms wordt het iets ingewikkelder, zoals de aanpassing van een theoretische verdeling aan een waargenomen frequentieverdeling bij een elementair wachttijdenprobleem (zie bijv. de figuur op blz. 231). En soms wordt het veel ingewikkelder, zoals het geval is bij de statistische bepaling van coëfficiënten van de vergelijkingen van een econometrisch macro-model. Dit laatste onderwerp zal nu ter sprake komen, maar we zullen ons daarbij beperkingen opleggen om de zaak redelijk eenvoudig te houden. We bepalen ons nl. tot een enkele vergelijking, die er bovendien niet al te gecompliceerd uitziet. Het betoog zal in drie ronden verlopen:

(1) Onmiddellijk hieronder beginnen we met de vraag welke de ingrediënten zijn van de procedure die uiteindelijk de resultaten zal opleveren. Het zal geen verwondering wekken, dat die ingrediënten o.a. uit zekere statistische gegevens bestaan.

(2) Vervolgens zullen we ons afvragen, hoe men m.b.v. die statistische gegevens het gevraagde verband kan berekenen.

(3) Is dit achter de rug, dan zullen we ons slechts gedeeltelijk bevredigd voelen omdat we nog zo weinig weten over de betrouwbaarheid van het resultaat. Dit geeft aanleiding tot de derde en laatste ronde, die op de waarschijnlijkheidsrekening is gebaseerd.

Algemeen kan men stellen, dat de statistische specificatie van economische relaties gebaseerd is op twee soorten van gegevens, nl. enerzijds hetgeen men al weet of meent te kunnen stellen omtrent de relatie in kwestie, anderzijds een aantal statistische gegevens. Wat de laatste betreft, deze kunnen verschil-

lende vormen aannemen. Men kan de beschikking hebben over huishoudrekeningen van gezinnen, die aan een huishoudrekeningenonderzoek hebben deelgenomen. De gegevens van bedrijven die hebben deelgenomen aan een investeringsenquête vormen een ander voorbeeld. Een derde (en in de practijk van het econometrisch onderzoek heel belangrijke) bron van informatie wordt gevormd door de tijdreeksen die door statistische bureaux zijn samengesteld. Stellen we ons bijv. voor, dat het gaat om het verband tussen consumptie per hoofd en inkomen per hoofd in Nederland sinds de Tweede Wereldoorlog; en laat het zo zijn dat er voor beide variabelen per jaar statistische gegevens van het C.B.S. beschikbaar zijn. Welnu, dan moet het aan de hand van deze gegevens mogelijk zijn iets naders over het gezochte verband te weten te komen.

Dit consumptie-inkomen voorbeeld is tevens geschikt om de betekenis van de andere bron van informatie te illustreren. Zoals reeds gezegd gaat het hier om hetgeen men a priori al weet of redelijkerwijs meent te kunnen stellen; en in dit verband ligt het voor de hand te vragen: wanneer men zich interesseert voor het gedrag in de loop van de tijd van de consumptie per hoofd, is het dan werkelijk aanvaardbaar om dit gedrag uitsluitend met behulp van het inkomen per hoofd te beschrijven? Het is natuurlijk wel eenvoudig om maar één variabele te introduceren en, toegegeven, eenvoud is een deugd op zichzelf, maar overdrijven we het niet? Het antwoord moet luiden, dat inderdaad aan bijzondere voorwaarden voldaan moet zijn opdat het inkomen per hoofd het werk alleen voor zijn rekening kan nemen. Laten we bijv. de bevolking van Nederland in tweeën opdelen: gezinnen van loontrekkers en overige gezinnen ('zelfstandigen'). Het is niet gezegd dat beide categorieën hetzelfde uitgaafpatroon hebben; denkbaar is immers, dat de zelfstandigen een groter deel van een gulden extra inkomen besparen, dus niet consumptief besteden, dan de loontrekkers. Het is ook denkbaar, dat de zelfstandigen met enige vertraging op veranderingen van hun inkomen reageren, terwijl de loontrekkers onmiddellijk reageren.[1] In dergelijke gevallen zal het aanbeveling verdienen de consumptie per hoofd niet te beschrijven m.b.v. het inkomen per hoofd van de gehele bevolking, maar bijv. met

1. Een voorbeeld hiervan is de consumptiefunctie van het kleine model van de Verenigde Staten, dat in Hoofdstuk 4 behandeld is.

het inkomen per hoofd van de loontrekkers in dit jaar;
het inkomen per hoofd van de zelfstandigen in dit jaar;
het inkomen per hoofd van de zelfstandigen in het vorige jaar.

Er is nog meer. Laten we gemakshalve veronderstellen, dat het inkomen per hoofd de enige belangrijke variabele is voor het gedrag in de loop van de tijd van de consumptie per hoofd. Dan nog is er een probleem, nl.: hoe luidt dit verband? Hoe is zijn wiskundige vorm? Het is mogelijk, dat de consumptie evenredig is met het inkomen; hij kan een lineaire functie zijn van het inkomen, of kwadratisch; er zijn nog talloze andere mogelijkheden. Zolang het gaat om het verband tussen twee variabelen kan men langs empirische weg het een en ander doen. Men kan nl. de waarnemingen in de vorm van punten in een diagram uitzetten en nagaan welke wiskundige vorm redelijk aansluit. Maar wanneer een derde variabele een rol speelt (zoals het geval is indien men de inkomens van loontrekkers en van zelfstandigen beide in een consumptievergelijking opvoert), dan wordt dit al erg omslachtig; en bij vier of meer variabelen is het onmogelijk. En aangezien het herhaaldelijk voorkomt, dat men drie variabelen gebruikt als verklaringsfactoren voor een vierde, is het dus noodzakelijk op een andere manier te werk te gaan. Deze manier komt hierop neer, dat men zich van tevoren bezint over de vorm van de relatie. Op basis hiervan voert men zekere berekeningen uit (waarover hieronder meer); en wanneer die gereed zijn bestaat er een mogelijkheid tot verificatie van de gekozen vorm.

De spelregels, die in de practijk van het econometrisch onderzoek worden gevolgd, komen op het volgende neer. Het object van onderzoek is de variatie van een of andere grootheid, verder de 'te verklaren variabele' te noemen. Dit kan zijn het gedrag van de totale consumptie gedurende een zekere periode; het kan ook zijn de variatie van het boterverbruik van gezin tot gezin in een huishoudrekeningenonderzoek, enz. Daartoe specificeert men (i) de variabelen die als verklaringsfactoren worden opgevoerd, verder 'verklarende variabelen' te noemen, en (ii) de mathematische vorm van de wijze, waarop de verklarende variabelen de te verklaren variabele beïnvloeden. Gaat het om de consumptie per hoofd, zoals in de twee vorige alinea's, dan impliceert de eerste stap dus dat we aangeven dat het inkomen per hoofd van de loontrekkers een verklarende variabele is, ook dat van de zelfstandigen, eventueel niet alleen

lopend maar ook een jaar vertraagd, enz. De tweede stap houdt in, dat we stellen dat de beïnvloeding lineair plaats heeft, of kwadratisch, of wat ook. Deze stappen zijn in beginsel gebaseerd op economisch-theoretische overwegingen. Het komt echter vaak voor, dat de economische theorie niet erg veel houvast geeft; dat geldt i.h.b. voor de wiskundige vorm van de relaties. In dat geval zullen ervaring, intuïtie en overwegingen van eenvoud een overheersende rol spelen bij de specificatie van het verband.

De eerste twee stappen geven wel aan, hoe de relatie eruit ziet, maar ze doen dat niet op een numerieke manier. Zo'n relatie is bijv. van het type $y = a + bx$, waarbij y en x consumptie resp. inkomen voorstellen en a en b vooralsnog onbekende coëfficiënten. Het is de volgende taak van de econometrist deze coëfficiënten met behulp van statistisch waarnemingsmateriaal door numerieke waarden te vervangen. Dat is dan de tweede ronde, die in het begin van dit hoofdstuk werd aangekondigd en waartoe we nu zullen overgaan.

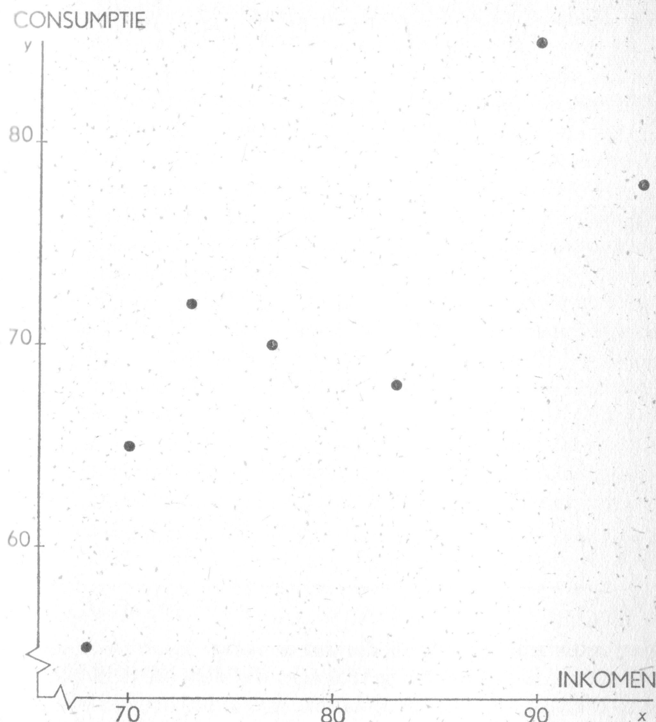
2. DE PUNTENWOLK

Neem weer het geval van twee variabelen: x (inkomen, verklarende variabele) en y (consumptie, te verklaren variabele). We gaan ervan uit, dat er een zeker aantal waarnemingen voor deze variabelen beschikbaar is. Bijvoorbeeld, in het eerst jaar 77 voor x en 70 voor y; in het tweede 83 voor x en 68 voor y, enz. Het waarnemingsmateriaal als geheel laat zich dan eenvoudig weergeven door een tweetal kolommen van getallen:

x	y
77	70
83	68
90	85
70	65
73	72
95	78
68	55

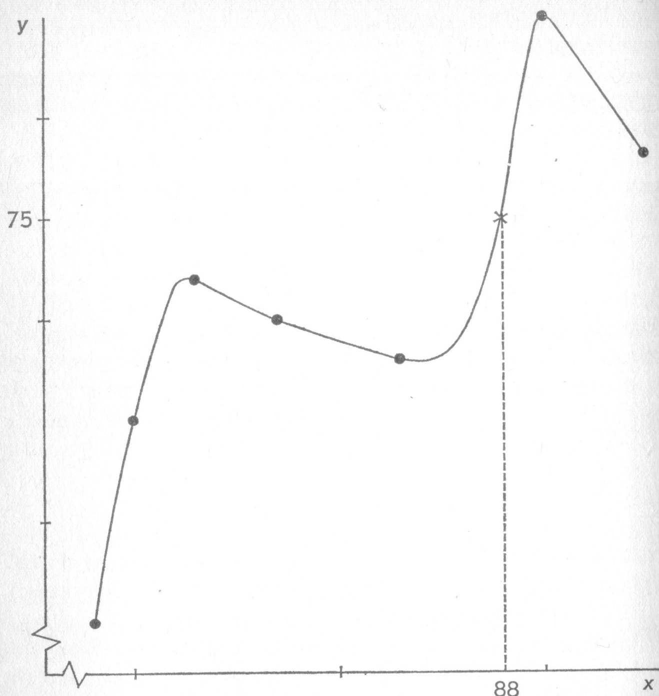
Deze getallenparen zet men vervolgens af in een diagram, waarvan de horizontale as bestemd is voor de verklarende variabele

(x) en de verticale voor de te verklaren variabele (y). Het eerste punt bevindt zich dus op een afstand 77 van de verticale as en een afstand 70 van de horizontale as; voor het tweede punt zijn deze afstanden 83 resp. 68, enz. We krijgen dan het beeld van een 'puntenwolk':



De vraag rijst nu, wat we met dit waarnemingsmateriaal moeten beginnen. Als de punten keurig in de pas op een rechte lijn hadden gelegen, was er geen enkel probleem geweest; we zouden die lijn kunnen trekken en hem als het verband tussen x en y kunnen interpreteren. Maar de punten liggen nu eenmaal niet op een rechte lijn; en bovendien, zoals de praktijk van dit soort werk leert, dat doen ze vrijwel nooit. De lezer zal wellicht tegenwerpen, dat het verband toch niet per se rechtlijnig behoeft te zijn; en we kunnen altijd een kromme lijn door die

punten aanbrengen. We gaan met hem mee en tekenen de volgende figuur:



De punten zijn dezelfde als die van de vorige figuur; we hebben op het oog een kromme getrokken, die door al deze punten gaat. (Dat kan natuurlijk op talloze manieren, hetgeen op zichzelf al een bezwaar tegen deze procedure is.) De vraag is: kunnen we deze kromme lijn interpreteren als het verband tussen *x* en *y*, dat we zoeken? Er is een overweging, die deze opvatting steunt. De kromme voldoet aan het waarnemingsmateriaal; zouden we hem zodanig tekenen, dat hij een of meer punten 'mist', dan voldoet hij niet aan de met deze punten corresponderende waarnemingen. En dit kan worden beschouwd als een strijdigheid van de theorie (de kromme) enerzijds en de feiten (de punten) anderzijds. Daar staan twee andere overwegingen tegenover. In de eerste plaats zien we, dat de getekende kromme

afwisselend daalt en stijgt. Dit betekent, dat als we die kromme als het 'ware' verband tussen *y* en *x* zouden interpreteren, een stijging van *x* soms tot een daling en soms tot een stijging van *y* aanleiding kan geven, afhankelijk van waar we precies zitten. Het behoeft geen betoog, dat dit niet erg plausibel is. Er zijn natuurlijk situaties van een tekenomslag van de beïnvloeding denkbaar, maar hier wordt het toch wel bont gemaakt. Een tweede bezwaar is dit: stel dat een nieuwe waarneming ter beschikking komt – zou ook die aan onze kromme lijn gehoorzamen? Preciezer: stel dat de *x*-waarde van de nieuwe waarneming 88 is; dan moet de bijbehorende *y*-waarde ongeveer 75 zijn wil het punt op de kromme liggen (zie de gestippelde verticale rechte). Geloven we dat dit altijd zo zal zijn? En als de *y*-waarde 70 is, wat dan te doen? Een nieuwe kromme tekenen? En dat keer op keer herhalen naarmate nieuwe waarnemingen ter beschikking komen?

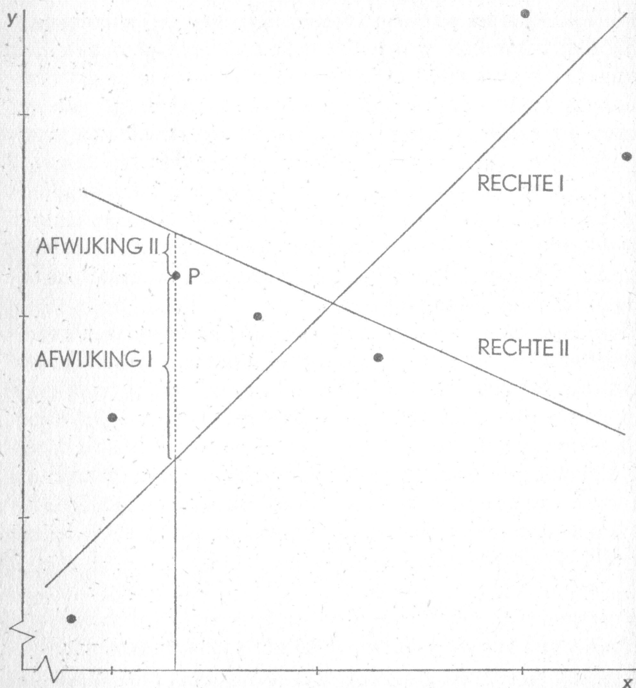
Deze bezwaren wegen zo zwaar dat de hier besproken kromme niet in tel is. Wat men in feite doet is het aanpassen van een rechte lijn (of een kromme lijn van een eenvoudig type) door de puntenwolk. Maar dit heeft belangrijke gevolgen. Het waarnemingsmateriaal voldoet niet, zoals in het geval van de kromme, aan de lijn die wij zullen construeren om het verband tussen *x* en *y* weer te geven. Ons verband zal niet 'volledig' zijn; het is het 'systematische verband' tussen *x* en *y* en dat betekent dat er afwijkingen zullen zijn.

3. AANPASSING VOLGENS KLEINSTE KWADRATEN

In hetgeen volgt zullen wij ervan uitgaan, dat het systematische verband tussen *x* en *y* lineair is, dus door een rechte lijn wordt voorgesteld. Daarmee hebben wij de keuze al aanzienlijk ingeperkt, maar het verband tussen de twee variabelen is nog lang niet numeriek gespecificeerd. Er zijn immers talloze rechte lijnen; en eerst als we één bepaalde rechte hebben geselecteerd ligt het verband numeriek vast.

Het probleem is dus: wat voor criterium moeten wij hanteren voor het eenduidig vastleggen van die rechte? Welnu, het is duidelijk in welke richting wij het moeten zoeken; we zullen het er immers over eens zijn, dat de afwijkingen liever klein dan groot moeten zijn. Nu geldt steeds, dat wanneer de puntenwolk ge-

geven is, *elke* rechte lijn voor *elk* punt van de puntenwolk een bepaalde afwijking heeft. Dit wordt hieronder geïllustreerd voor een van de punten (P) van de reeds eerder geschetste wolk en een tweetal alternatieve rechten.



Merk op, dat we de afwijkingen verticaal meten. Dat ligt in de rede, omdat het gedrag van de verticaal afgezette variabele y (consumptie in ons voorbeeld) het object van onderzoek is. Onze vraagstelling kan in deze vorm worden gegoten: wat kan gezegd worden over de consumptie (y) als gegeven is dat het inkomen (x) een zekere waarde aanneemt? Welnu, dan sporen we die gegeven x-waarde langs de horizontale as op, gaan verticaal omhoog om te constateren dat de y-waarde volgens de rechte lijn daar-en-daar moet liggen, en vinden tenslotte dat de waargenomen y-waarde hoger of lager ligt.

In de getekende figuur is het kennelijk zo, dat de Rechte I het

er beter afbrengt dan Rechte II. Toch zien we, dat de laatste t.o.v. het hier in bijzonder beschouwde punt P een geringere afwijking heeft dan Rechte I. Blijkbaar is de prestatie van zo'n rechte t.o.v. een enkel punt in het totaalbeeld van ondergeschikt belang. We kunnen in beginsel een rechte door twee van de punten tekenen, dan zijn de bijbehorende afwijkingen zelfs gereduceerd tot nul. Dat kunnen we op verscheidene manieren doen, nl. door de punten op alle mogelijke manieren in paren te nemen en door dan zo'n paar met een rechte lijn te verbinden. Maar aldus doende verwaarloost men alle overige punten.

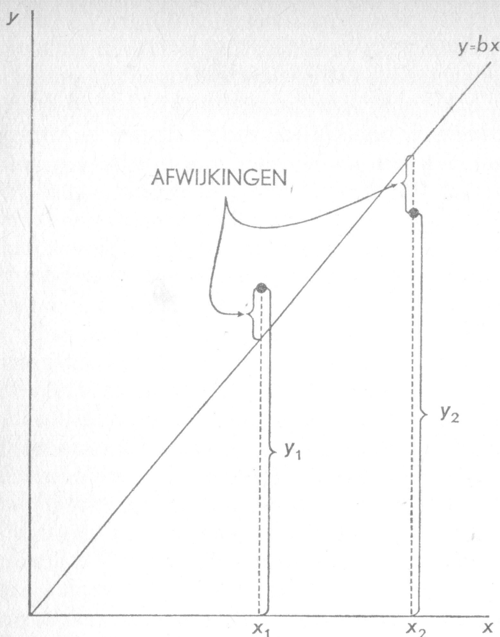
Een 'redelijke' manier van aanpassen houdt met alle punten tegelijkertijd rekening. Verreweg het bekendst is 'de methode der kleinste kwadraten', die het volgende criterium hanteert: neem van elke afwijking het kwadraat, tel al die kwadraten op, en zoek dan van alle denkbare rechte lijnen *die* op waarvoor de kwadratensom zo klein mogelijk is. Deze methode gaat terug tot de mathematicus Gauss uit het begin van de negentiende eeuw en is sindsdien op grote schaal en op talrijke terreinen toegepast.

De wijze, waarop wij het criterium van kleinste kwadraten zojuist formuleerden, suggereert wellicht dat we de kwadratensommen van de afwijkingen voor een groot aantal rechte lijnen ook in feite zouden moeten uitrekenen. Dit zou bepaald onpractisch zijn; gelukkig is het ook niet nodig. Wij zullen dit laten zien voor het eenvoudigste geval. Daartoe gaan we uit van de veronderstelling, dat het bekend is dat de rechte door de oorsprong moet lopen; anders gezegd, het op te sporen verband is er een van evenredigheid, $y = bx$, waarbij b de coëfficiënt is die numeriek moet worden vastgelegd. Voorts gaan we in eerste aanleg ervan uit, dat er maar twee waarnemingen zijn. Dit geval is weergegeven in de figuur op de volgende bladzijde, waar de twee punten corresponderen met de twee waarnemingen. Hierbij stellen x_1 en y_1 de coördinaten van de eerste waarneming voor; analoog x_2 en y_2 voor de tweede.

Neem nu een willekeurige rechte door de oorsprong, waarvan de vergelijking dus luidt $y = bx$. Vullen we $x = x_1$ in, dan vinden we bx_1 als y-waarde. De waargenomen waarde is echter y_1, dus is de afwijking $y_1 - bx_1$. Op dezelfde manier vinden we voor de tweede waarneming een afwijking $y_2 - bx_2$. De som van de kwadraten van de afwijkingen is dus

$$(y_1 - bx_1)^2 + (y_2 - bx_2)^2;$$

en ons criterium zegt, dat we de rechte lijn zodanig moeten



selecteren, dat deze uitdrukking zo klein mogelijk is. Maar we hebben zojuist gezien, dat de vergelijking van de rechte luidt $y = bx$; de rechte wordt dus volledig bepaald door de coëfficiënt b. Anders gezegd, we zullen ons probleem opgelost hebben zodra het criterium een eenduidige waarde voor b heeft opgeleverd. Daartoe werken we de kwadratensom als volgt uit:

$$(y_1^2 - 2bx_1y_1 + b^2x_1^2) + (y_2^2 - 2bx_2y_2 + b^2x_2^2)$$
$$= (y_1^2 + y_2^2) - 2b(x_1y_1 + x_2y_2) + b^2(x_1^2 + x_2^2).$$

Aldus hebben we de kwadratensom opgesplitst in drie termen naar opklimmende macht van b: de eerste term, $y_1^2 + y_2^2$, bevat b niet; de tweede, $-2b(x_1y_1 + x_2y_2)$, is kennelijk van de eerste graad in b; de derde, $b^2(x_1^2 + x_2^2)$, is van de tweede graad. We kunnen de kwadratensom op nog een andere manier schrijven, nl. als de som van

(1) $$y_1^2 + y_2^2 - \frac{(x_1y_1 + x_2y_2)^2}{(x_1^2 + x_2^2)}$$

en

$$(2)\quad (x_1^2 + x_2^2)\left[b - \frac{x_1y_1 + x_2y_2}{x_1^2 + x_2^2} \right]^2.$$

Dat dit inderdaad op hetzelfde neerkomt vinden we gemakkelijk door term (2) uit te werken en dan op te tellen bij term (1).

Nu zien we onmiddellijk, dat term (1) de coëfficiënt b niet bevat. Dit betekent, dat hoe we b ook kiezen (hoe we de rechte lijn ook tekenen), deze term altijd dezelfde zal zijn. Anders gezegd, wanneer het erom gaat de kwadratensom van de afwijkingen zo klein mogelijk te maken, dan kunnen we het deel van deze kwadratensom dat door term (1) wordt gerepresenteerd buiten beschouwing laten, want daar valt toch niets aan te doen. Wij kunnen ons er dus toe bepalen het restant, d.w.z. term (2) zo klein mogelijk te maken. Nu is deze term gelijk aan $x_1^2 + x_2^2$ vermenigvuldigd met het kwadraat van de uitdrukking tussen vierkante haken. Aangezien de vermenigvuldigingsfactor $x_1^2 + x_2^2$ altijd positief is (of nul, maar dat geval is niet interessant), bereiken we de minimale waarde door de uitdrukking tussen vierkante haken nul te maken. Dit betekent, dat b gespecificeerd dient te worden volgens de formule

$$b = \frac{x_1y_1 + x_2y_2}{x_1^2 + x_2^2};$$

en aangezien deze keuze van b de term (2) nul maakt, wordt de kwadratensom van de afwijkingen dus gegeven door term (1). Deze term stelt de minimale waarde voor, die voor deze kwadratensom bereikbaar is. Wanneer niet twee waarnemingen maar een willekeurig aantal beschikbaar is, dan vindt men op analoge wijze na een beetje meer rekenwerk:

$$b = \frac{x_1y_1 + x_2y_2 + x_3y_3 + \ldots}{x_1^2 + x_2^2 + x_3^2 + \ldots},$$

in woorden: we gaan uit van alle beschikbare waarnemingsparen (x_1, y_1), (x_2, y_2), ...; we rekenen de producten x_1y_1, x_2y_2, ... per waarnemingspaar uit en tellen die op en plaatsen de som in de teller; we rekenen de kwadraten x_1^2, x_2^2, ... van de waarden aangenomen door de verklarende variabele uit en tellen die op en plaatsen de som in de noemer; en dan delen. Voor ons numerieke voorbeeld op blz. 310 vinden we de volgende teller:

$$77 \times 70 + 83 \times 68 + \ldots + 68 \times 55 = 39640$$

en de volgende noemer:

$$77^2 + 83^2 + \ldots + 68^2 = 44796,$$

zodat $b = 39640 / 44796 = 0{,}885$. De conclusie luidt dus: gegeven het ter beschikking staande waarnemingsmateriaal en onder de veronderstelling van een evenredigheidsverband, varieert de consumptie met het inkomen volgens de formule $y = 0{,}885x$, d.w.z. de consumptie bedraagt $88^1/_2\%$ van het inkomen – althans behoudens de onvermijdelijke afwijkingen (welker kwadratensom we geminimeerd hebben).

De vergelijking $y = bx$ met b zoals hier gespecificeerd wordt een *regressievergelijking* genoemd en wel een regressievergelijking volgens kleinste kwadraten. De rechte lijn die de grafische representatie is van deze vergelijking heet een *regressielijn* en b heet de *regressiecoëfficiënt*. De term 'regressie' is afkomstig van de negentiende-eeuwse Engelse statisticus Galton. Hij voerde dit soort berekeningen uit om het verband tussen de lengten van vaders en zoons te bepalen en concludeerde, dat lange vaders gemiddeld genomen ook wel lange zoons hebben maar niet zo lang als zij zelf zijn; en analoog, dat kleine vaders gemiddeld genomen kleine zoons hebben maar niet zo klein. Aldus treedt er een terugkeer, een regressie naar de gemiddelde lengte op in successieve generaties. Maar tegenwoordig is 'regressieanalyse' een op vele terreinen toegepaste techniek; de bijzondere interpretatie van Galton is daarom voor ons van weinig belang (en bovendien hebben de experts van vandaag niet veel waardering voor die bijzondere interpretatie).

Wij hebben ons hier bepaald tot het eenvoudigste geval, nl. dat van een rechte door de oorsprong. Wanneer die rechte niet door de oorsprong gaat is zijn vergelijking van het type $y = a + bx$, zodat we dus twee coëfficiënten moeten bepalen, a en b. Wanneer we met twee of meer verklarende variabelen werken, wordt dit aantal nog groter. De zaak wordt daardoor ingewikkelder, maar er komen geen nieuwe principiële problemen bij. Het criterium blijft het minimeren van de som van de kwadraten van de afwijkingen tussen enerzijds de waargenomen waarden van de te verklaren variabele en anderzijds de corresponderende waarden volgens de regressievergelijking. Het resultaat is dan, dat *alle* coëfficiënten *gelijktijdig* door dit criterium worden bepaald.

4. DE GEDACHTENGANG VAN HET STATISTISCH SCHATTEN

Wij zullen ons nu gaan bezig houden met het vraagstuk van de betrouwbaarheid van het verkregen resultaat, dus de derde ronde van het in het begin van dit hoofdstuk aangekondigde programma. Weliswaar is het zo, dat de kleinste-kwadratenprocedure tot een eenduidige uitkomst leidt, gegeven het waarnemingsmateriaal; wij hebben immers gezien hoe men de coëfficiënt b met behulp van dit materiaal kan berekenen. Maar aan de andere kant zal het bijv. intuïtief duidelijk zijn, dat een omvangrijk waarnemingsmateriaal ons in beginsel in staat stelt een betere uitspraak te doen omtrent de beïnvloeding van y door x dan een hoeveelheid materiaal van geringere omvang. De verdere bladzijden van dit hoofdstuk zullen aan dit onderwerp gewijd zijn. Het gaat hier om de zgn. statistische schattingsprocedures. Om misverstand te vermijden vermelden we nu al, dat het woord 'schatting' hier een statistisch-technische betekenis zal hebben. De termen 'schatting' en 'raming' worden in het dagelijks leven gebruikt voor benaderingen van iets dat niet met volmaakte nauwkeurigheid valt vast te stellen; en men hoopt dat zulke benaderingen niet te ver verwijderd zullen zijn van de exacte waarde. In de meeste gevallen wordt er echter weinig gezegd over de kwaliteit van de benadering, voornamelijk omdat het zo moeilijk is dit op goede gronden te doen. Maar in de statistische schattingstheorie staat het onderwerp van de betrouwbaarheid juist centraal. Daartoe maakt men gebruik van de waarschijnlijkheidsrekening.

Nu is de regressieanalyse een onderwerp uit de statistiek, dat zich bezig houdt met twee variabelen (of meer, nl. als er verscheidene verklarende variabelen zijn). We kunnen de gedachtengang van het statistisch schatten op eenvoudiger wijze illustreren door met slechts één variabele te werken. Wij maken dus nu een omweg alvorens tot de regressieanalyse terug te keren en daartoe bezien we opnieuw de 50 gezinnen, wier uitgaafbedragen op blz. 173 zijn afgedrukt bij de introductie van de normale verdeling in Hoofdstuk 6. In een later stadium hebben we 150 gezinnen opgevoerd en in een nog later stadium lieten we dit aantal onbeperkt oplopen. Alle gezinnen, bestaande uit man, vrouw en twee kinderen, met een inkomen van f 1000,— per maand.

Bepalen wij ons eerst tot die 50 gezinnen. Wanneer wij hun

uitgaafpatroon met een enkel getal willen karakteriseren, dan ligt het voor de hand om bijv. het rekenkundig gemiddelde te nemen. Algebraïsch: als $R_1, R_2, \ldots, R_{50}$ de 50 uitgaafbedragen zijn, dan is hun rekenkundig gemiddelde m:

$$m = \frac{R_1 + R_2 + \ldots + R_{50}}{50}.$$

En numeriek voor de op blz. 173 gegeven waarden:

$$m = \frac{716 + 937 + \ldots + 839}{50} = \frac{41048}{50} = 820{,}96.$$

De conclusie luidt dus, dat het gemiddelde uitgaafbedrag van *die* vijftig gezinnen gelijk is aan (afgerond) f 821. Een conclusie, die interessant is wanneer we ons voor die 50 gezinnen bijzonder interesseren. Maar al gauw rijst dan de vraag, of het zo vaak zal voorkomen, dat men zich voor het uitgaafpatroon van een dergelijk vijftigtal bijzonder interesseert. Ons antwoord luidt: neen, maar het zal wèl veel vaker voorkomen, dat men interesse heeft voor bijv. het gemiddelde uitgaafbedrag van *alle* Nederlandse (of alle Zuidhollandse of alle Rotterdamse) gezinnen bestaande uit man, vrouw en twee kinderen met een inkomen van f 1000 per maand. Deze overweging nu ligt ten grondslag aan een belangrijk onderscheid in de statistische theorie. We hebben enerzijds de 50 gezinnen, die ons gegevens verstrekten omtrent hun uitgaven. Anderzijds is er de veel grotere groep van alle gezinnen in dezelfde omstandigheden (man, vrouw, twee kinderen, enz.), en die bevat de eerste groep als een klein onderdeel. We noemen de eerste groep een *steekproef*, die getrokken is uit de tweede, die op zijn beurt *populatie* heet. Preciezer gezegd, in het onderhavige geval bestaat de populatie uit de uitgaafbedragen van alle gezinnen bestaande uit man, vrouw en twee kinderen met een inkomen van f 1000 per maand; en de steekproef is eenvoudig dat onderdeel van die populatie waarvoor die uitgaafbedragen zijn waargenomen. Dit onderscheid kan op grote schaal en op talrijke terreinen gemaakt worden. Bijvoorbeeld, een importeur van mengvoeder interesseert zich voor het eiwitgehalte van een partij, die hij zojuist heeft ontvangen. Daartoe laat hij monsters nemen uit een paar zakken, die vervolgens op het laboratorium worden onderzocht. Dan hebben we enerzijds het eiwitgehalte van de monsters (de steekproef), anderzijds het eiwitgehalte van

de partij als geheel (de populatie). Zo zijn er talloze andere voorbeelden, waaronder een aantal van meer abstracte aard; men kan bijv. het resultaat van tienmaal een munt opwerpen (zegge 6 maal kruis en 4 maal munt) beschouwen als een steekproef uit de populatie bestaande uit alle denkbare werpresultaten.

De statistische schattingstheorie dankt zijn bestaan aan het probleem van de onvolledige kennis. Onze kennis beperkt zich immers tot de steekproef, maar daarvoor interesseren we ons eigenlijk niet; we beschouwen die steekproef slechts als een hulpmiddel om te komen tot een grotere kennis omtrent de populatie. Zo gaat het de importeur van mengvoerder om het eiwitgehalte van de partij als geheel; het gehalte van de monsters is voor hem alleen maar van belang voorzover dit hem licht kan verschaffen over het totale gehalte. Analoog zijn de uitgaafbedragen van onze 50 gezinnen slechts interessant voorzover ze de bijbehorende populatie in een helderder licht brengen. Nu zal de lezer wellicht tegenwerpen, dat dit laatste probleem al lang voor ons is opgelost; op blz. 177 is immers met zoveel woorden gesteld, dat die uitgaven normaal verdeeld zijn met verwachting $\mu = 800$ en standaarddeviatie $\sigma = 100$. Dat is juist, en de populatie is daarmee volledig gedefinieerd. Maar het is duidelijk, dat dit niet de typische situatie is. De gegevens zijn door ons op deze manier geconstrueerd; het statistisch onderzoek houdt zich daarentegen bezig met het geval van een niet volledig bekende populatie.

Teneinde het probleem nog wat concreter te maken zullen wij ons verder bepalen tot de gemiddelde uitgaven. Wij beschikken over een steekproef van 50; zoals we hierboven hebben gezien is deze gekenmerkt door een gemiddeld uitgaafbedrag van f 821. Stel nu eens, dat wij deze uitkomst wensen te gebruiken als een benadering van het gemiddelde uitgaafbedrag van alle gezinnen uit de populatie. Welnu, de kwaliteit van deze benadering is in dit geval wel heel eenvoudig te beoordelen, want het populatiegemiddelde is niets anders dan de verwachting μ, dus f 800. De conclusie is dus dat steekproefgemiddelde en populatiegemiddelde van elkaar afwijken, dus dat de benadering niet volmaakt is; voorts, dat de afwijking f 21 bedraagt. Of men die afwijking groot of klein moet vinden is een vraag, die slechts beantwoord kan worden, wanneer men concreet aangeeft wat met de uitkomst wordt gedaan. Belangrijker is echter, dat we ons goed realiseren dat we ons hier in een uit-

zonderlijke situatie bevinden. We kennen de verwachting en kunnen dus de benaderingsfout berekenen. In het algemeen zullen we die verwachting niet kennen, zodat we naar andere middelen moeten uitzien om uit te maken in hoeverre het steekproefgemiddelde een redelijke benadering van het (onbekende) populatiegemiddelde is.

Laten we ons nu voorstellen, dat we een andere groep van 50 gezinnen uit dezelfde populatie selecteren, een tweede steekproef dus. Zal het gemiddelde uitgaafbedrag van de tweede steekproef wel gelijk zijn aan het populatiegemiddelde? Zal het tweede steekproefgemiddelde gelijk zijn aan het eerste steekproefgemiddelde? Uiteraard moet het antwoord op beide vragen luiden: het kan, maar het zou wel toevallig zijn. Laten we dus aannemen, dat het tweede steekproefgemiddelde bijv. f 803 bedraagt. Zo kunnen we doorgaan met een derde steekproef, een vierde, enz.; iedere keer berekenen we het steekproefgemiddelde en meestal vinden we een andere waarde. Zouden we deze procedure 50 maal voortzetten, zodat we dus 50 keer een steekproef van 50 gezinnen uit dezelfde populatie trekken, dan vinden we ook 50 gemiddelden die op dezelfde manier als op blz. 173 gerangschikt kunnen worden. Er is echter een belangrijk verschil. Eerst hadden we slechts 50 gezinnen en individuele uitgaafbedragen (per gezin); nu hebben we $50 \times 50 = 2500$ gezinnen, gerangschikt in groepen van 50, en de genoteerde uitgaafbedragen zijn geen individuele uitgaven maar gemiddelden per groep. Hoe zal die nieuwe lijst er uitzien? Het ligt in de rede, dat die gemiddelden met geringere spreiding rond het populatiegemiddelde (f 800) zullen zijn gestrooid dan geldt voor de lijst op blz. 173. Er is al enige indicatie in deze richting. De lijst op blz. 173 bevat immers een aantal gevallen van boven f 1000 en beneden f 600, die dus meer dan f 200 van het populatiegemiddelde verwijderd zijn; daarentegen bleek het eerste steekproefgemiddelde slechts f 21 van het populatiegemiddelde verwijderd te zijn.

Toch is het bepaald niet steeds zo, dat een steekproefgemiddelde noodzakelijkerwijs dicht bij het populatiegemiddelde ligt. Neem bijv. het geval van een steekproef van gezinnen, die geselecteerd zijn uit de groep van hen die ontsparen. Aangezien het inkomen op f 1000 ligt betekent dit, dat hun uitgaven alle minstens f 200 boven het niveau van f 800 liggen. Het gemiddelde uitgaafbedrag van een dergelijke steekproef ligt dan ook

meer dan f 200 boven het populatiegemiddelde. Dit voorbeeld lijkt wellicht triviaal, maar het illustreert op duidelijke wijze, dat de *trekkingsprocedure* van groot belang is. Een andere procedure is deze: selecteer die gezinnen voor de steekproef, wier uitgaafbedrag dicht ligt bij het populatiegemiddelde. Het lijdt geen twijfel dat deze procedure beter is voor de benadering van het populatiegemiddelde. Echter, hij brengt ons niet veel verder, omdat de uitvoering ervan vereist dat we weten wanneer een uitgaafbedrag dicht bij het populatiegemiddelde ligt. En het gaat juist om de kennis van dat gemiddelde!

De procedure die men pleegt te volgen komt neer op het bestrijden van vuur met vuur. We hebben te kampen met de variabiliteit van de gezinsuitgaven zoals die wordt afgebeeld door de normale verdeling van blz. 176. Deze variabiliteit is het vuur dat bestreden moet worden in de zin dat we enige orde op zaken willen stellen; d.w.z. 'orde' door althans het gemiddelde van de groep als geheel (μ) op adequate wijze te specificeren. Het middel om tot dit doel te geraken is het trekken van een steekproef, hetgeen op een speciale manier geschiedt. Niet, uiteraard, door uitsluitend gezinnen met hoge of uitsluitend gezinnen met lage uitgaven te selecteren. Maar, paradoxaal genoeg, door helemaal niet te selecteren; door de steekproef (met een fraaie term van wijlen de Amsterdamse hoogleraar Van Dantzig) *aselect* te trekken. Dat wil zeggen, door bij de 'verkiezing' van de eerste kandidaat voor de steekproef alle gezinnen van de populatie gelijke kans te geven gekozen te worden, door vervolgens bij de verkiezing van de tweede kandidaat aan alle niet in eerste instantie gekozen kandidaten gelijke waarschijnlijkheid te geven om gekozen te worden, enz. Merk op, dat de procedure van aselect trekken alleen maar iets specificeert over de *wijze* van trekken, niet over de vraag welke getalswaarden worden getrokken. Het is dus denkbaar, dat bij aselect trekken uit de hier besproken populatie vijfmaal achtereen een uitgaafbedrag boven f 1000 wordt getrokken (al zal het niet vaak voorkomen). In feite redeneert de aselecte procedure aldus: Gij, populatie, veroorlooft U de luxe ons zekerheid te onthouden ten aanzien van hetgeen de consument besteedt; welnu, dan zal ik U zekerheid onthouden ten aanzien van de groep van consumenten die ik daarover zal ondervragen – ik zal allen gelijke kans geven ondervraagd te worden. En in deze zin bestrijdt de procedure vuur met vuur.

Wat valt er te zeggen, wanneer de trekkingsprocedure van het aselecte type is? Om deze vraag te beantwoorden zullen wij ervan uitgaan, dat de op blz. 173 weergegeven uitgaafbedragen het resultaat zijn van aselect trekken. Nu is het van groot belang te doorzien, dat er eigenlijk sprake is van twee stadia en dat de afgedrukte getallen het resultaat vormen van het tweede stadium. Immers, aselect trekken betekent: we geven alle gezinnen van de populatie gelijke kans om in de steekproef te worden opgenomen. Dat is het eerste stadium. Vervolgens trekken we de steekproef, en dat is het tweede. Is eenmaal getrokken, dan liggen de uitkomsten numeriek vast, zie de lijst op blz. 173, en het enige interessante dat we dan nog kunnen zeggen is dat die uitkomsten zijn zoals ze zijn. Uiteraard kunnen we met die uitkomsten nog de nodige berekeningen uitvoeren, daar zijn ze voor en dat zullen we ook doen; maar het resultaat van het tweede stadium (de uitkomsten in numerieke vorm) is niet hetgeen waarmee de statistische theorie zich in eerste instantie bezig houdt.

Het eerste stadium is de essentie van aselect trekken: alle gezinnen van de populatie (of algemener, alle 'elementen' van de populatie) hebben gelijke kans getrokken te worden. Nu is het duidelijk, dat zolang wij nog in dit stadium verkeren – zolang er in feite nog geen element is getrokken – de steekproef alleen maar uit kansvariabelen bestaat. Bijvoorbeeld, laten wij terugkeren naar blz. 320, waar het steekproefgemiddelde voor onze 50 gezinnen werd afgeleid. We gaven daar de uitgaafbedragen weer met R_1, R_2, . . ., R_{50}, die vervolgens werden gespecificeerd als f 716, f 937, . . ., f 839. Maar deze numerieke specificatie is een zaak van het tweede stadium. Zijn we daar nog niet aan toe, dan is R_1 alleen maar een symbolische schrijfwijze voor het uitgaafbedrag van het eerste gezin dat in de steekproef *zal* worden opgenomen; analoog R_2 voor dat van het tweede 'toekomstige' gezin, enz. En inderdaad, zolang er alleen maar kansuitspraken worden gedaan, kunnen we R_1, R_2, . . . slechts als kansvariabelen interpreteren. Maar we kunnen nog iets meer zeggen; we kunnen nl. hun verdeling aangeven op grond van het feit, dat uitdrukkelijk gestipuleerd is, dat alle gezinnen *gelijke* kans op opname in de steekproef hebben. Men kan dit het gemakkelijkst inzien door de zaak precies om te draaien.

Stel nl., dat wij de steekproef zodanig zouden trekken, dat gezinnen met een uitgaafbedrag beneden f 750 een 100 maal zo kleine kans hebben als die boven f 750. Hoe zou de verdeling van $R_1, \ldots, R_{50}$ er dan uit zien? Welnu, de dichtheid van die verdeling zal een stuk lager zijn dan de normale dichtheid van blz. 176 voor bedragen beneden f 750, voor bedragen daarboven zal hij daarentegen hoger liggen. Maar dat is nu juist het punt: als we alle elementen van de populatie *gelijke* kans geven getrokken te worden, dan hebben de kansvariabelen die tezamen de steekproef vormen alle identiek dezelfde verdeling als die van de populatie. En dit geldt algemeen. Aselect trekken betekent – juist vanwege de gelijke kansen – dat we trekken conform de populatieverdeling.

We zijn nu zover, dat we tot resultaten kunnen komen. Laten we daartoe uitgaan van de veronderstelling, dat bekend is dat de uitgaafbedragen normaal verdeeld zijn maar dat daarentegen μ en σ^2 (populatiegemiddelde en variantie) onbekend zijn. Ons doel is μ te leren kennen, althans te benaderen. Daartoe gaan we een aselecte steekproef van 50 gezinnen trekken en we beschouwen het steekproefgemiddelde, dat we in de volgende vorm opschrijven:

$$m = \frac{1}{50}R_1 + \frac{1}{50}R_2 + \ldots + \frac{1}{50}R_{50}.$$

Nu zijn $R_1, R_2, \ldots$ kansvariabelen zolang in feite nog niet is getrokken. Maar dan is ook m een kansvariabele, nl. een lineaire combinatie van de 'oorspronkelijke' kansvariabelen $R_1, R_2, \ldots$ Wij weten (zie blz. 171), dat de verwachting van een lineaire combinatie van kansvariabelen gelijk is aan dezelfde lineaire combinatie van de verwachtingen van die kansvariabelen; bovendien weten we, dat $R_1, R_2, \ldots$ alle verwachting μ hebben, omdat zij alle onderhevig zijn aan dezelfde normale verdeling met verwachting μ. Dus is de verwachting van m

$$\frac{1}{50}\mu + \frac{1}{50}\mu + \ldots + \frac{1}{50}\mu = 50 \times \frac{1}{50}\mu = \mu,$$

d.w.z. de *verwachting* van het steekproefgemiddelde is in geval van aselect trekken gelijk aan het populatiegemiddelde. Anders gezegd, wanneer onze steekproefprocedure neerkomt op aselect trekken, dan zal het resultaat in de vorm van het steekproefgemiddelde nooit gefixeerd zijn maar onderhevig zijn aan

een verdeling; soms zal de in feite gerealiseerde waarde van het steekproefgemiddelde boven, soms beneden het populatiegemiddelde terecht komen. Maar er is één ding, dat we met zekerheid kunnen zeggen: de aselecte procedure garandeert, dat 'gemiddeld genomen' het steekproefgemiddelde met het populatiegemiddelde samenvalt. Preciezer gezegd: wanneer we voor iedere numerieke realisatie van de aselecte procedure het verschil zouden nemen van het resulterende steekproefgemiddelde en het populatiegemiddelde, dan is dit verschil in het algemeen niet nul; maar wel heffen dergelijke verschillen elkaar op in de zin dat ze soms positief en soms negatief zijn en gemiddeld gelijk aan nul. In deze situatie spreekt men van een *zuivere schatting*.

Het wordt nu tijd om tot de regressieanalyse terug te keren.

6. SCHATTING IN DE REGRESSIEANALYSE

Stel dat wij over de huishoudrekeningen van een drietal gezinnen beschikken. In tegenstelling tot het geval van de voorgaande bladzijden nemen wij nu niet aan dat hun inkomens alle gelijk zijn aan f 1000 per maand; wij gaan ervan uit, dat die inkomens verschillend zijn, nl. f 1000, f 1500 en f 2500. Laat voorts gegeven zijn, dat de respectieve uitgaafbedragen als volgt zijn (ook in guldens per maand):

716 1268 2011.

Het is onze taak de samenhang tussen inkomen en uitgaven aan de hand van deze gegevens op te sporen. Welnu, daarvoor hebben wij een procedure ontwikkeld. Laten we x schrijven voor inkomen (de verklarende variabele) en y voor uitgaven (de te verklaren variabele); laten we er gemakshalve van uitgaan, dat het gepostuleerde verband neerkomt op evenredigheid, dus $y = bx$. Dan hebben we de volgende waarnemingsparen:

$$x_1 = 1000 \quad y_1 = 716$$
$$x_2 = 1500 \quad y_2 = 1268$$
$$x_3 = 2500 \quad y_3 = 2011$$

Passen we de formule van blz. 317 toe, dan vinden we voor de regressiecoëfficiënt:

$$b = \frac{x_1y_1 + x_2y_2 + x_3y_3}{x_1^2 + x_2^2 + x_3^2}$$

$$= \frac{1000 \times 716 + 1500 \times 1268 + 2500 \times 2011}{(1000)^2 + (1500)^2 + (2500)^2} = 0{,}8048.$$

De conclusie luidt dus, dat de uitgaven uit het bijbehorende inkomen kunnen worden afgeleid door vermenigvuldiging met de coëfficiënt 0,8048.

Nu zal het uit het voorgaande duidelijk zijn, dat over de stabiliteit en de betrouwbaarheid van dit numerieke resultaat het laatste woord nog niet is gezegd. We hebben immers maar één geval van een gezin met f 1000 inkomen, één geval met f 1500 inkomen en één geval met f 2500. Het heeft instructieve waarde ons waarnemingsmateriaal nog iets uit te breiden. We nemen dus een tweede gezin met f 1000 inkomen, ook een tweede met f 1500 inkomen en analoog voor f 2500. Laten de respectieve uitgaven als volgt zijn (wederom in guldens per maand):

937 1123 1911.

Ook voor dit geval kunnen we de samenhang tussen uitgaven en inkomen vaststellen en wel op dezelfde manier. Dit leidt tot de volgende waarde van b:

$$b = \frac{1000 \times 937 + 1500 \times 1123 + 2500 \times 1911}{(1000)^2 + (1500)^2 + (2500)^2} = 0{,}7788.$$

De conclusie luidt nu, dat de uitgaven uit het bijbehorende inkomen worden verkregen door vermenigvuldiging met 0,7788.

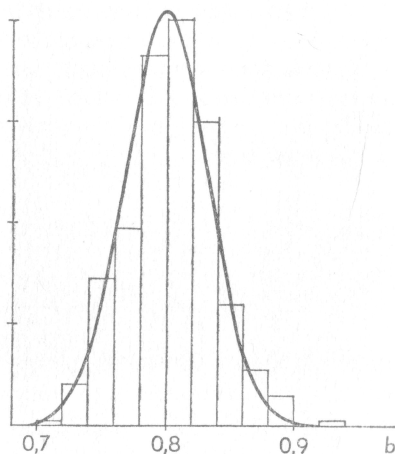
Kennelijk is er een tegenspraak tussen de numerieke conclusie, die gebaseerd is op het eerste waarnemingsmateriaal, en die van het tweede materiaal. Dit probleem wordt in de onderstaande tabel nader uitgewerkt door in totaal 50 groepen van waarnemingen te beschouwen. Zij zijn genummerd van 1 t/m 50, zie de voorkolom. In de drie kolommen daarna zijn de uit-

Steekproef nummer	Uitgaven bij een inkomen van f 1000	f 1500	f 2500	Regressiecoëfficiënt (b)
1	716	1268	2011	0,8048
2	937	1123	1911	0,7788
3	782	1061	2099	0,8022
4	835	1009	2050	0,7867
5	1082	1205	1905	0,8055
6	1012	1332	2081	0,8645
7	701	1133	1972	0,7716

Steekproef nummer	Uitgaven bij een inkomen van f 1000	f 1500	f 2500	Regressiecoëfficiënt (*b*)
8	834	1119	1836	0,7476
9	864	1296	1856	0,7840
10	811	1176	2001	0,7976
11	798	1134	2024	0,7957
12	714	1308	2046	0,8201
13	801	1227	2002	0,8049
14	797	1246	1791	0,7519
15	759	1123	2208	0,8383
16	862	1200	1920	0,7855
17	883	1301	2027	0,8318
18	722	1108	2021	0,7828
19	803	1298	1867	0,7808
20	570	1224	2170	0,8243
21	878	1301	1908	0,7999
22	866	1349	2173	0,8760
23	766	1316	2066	0,8321
24	916	1248	1831	0,7753
25	876	1240	2213	0,8704
26	700	1188	1780	0,7297
27	768	1115	2157	0,8245
28	818	1154	1815	0,7459
29	808	1235	1776	0,7474
30	998	1223	1985	0,8205
31	891	1220	1955	0,8009
32	768	1183	1997	0,7932
33	623	1212	2101	0,8098
34	816	1052	2119	0,8096
35	879	1252	1897	0,7894
36	820	1356	2035	0,8359
37	931	1304	1918	0,8086
38	835	1182	2176	0,8472
39	874	1174	1932	0,7858
40	827	1300	1912	0,7955
41	939	1149	1986	0,8029
42	786	1205	2077	0,8196
43	724	1131	1839	0,7387
44	837	1142	1933	0,7771
45	572	1131	2058	0,7804
46	836	1059	1894	0,7536
47	937	1224	2051	0,8316
48	904	1241	1976	0,8111
49	733	1215	1840	0,7532
50	839	1176	2029	0,8079

gaafbedragen gespecificeerd behorende bij de drie hier beschouwde inkomens: f 1000, f 1500 en f 2500. Zo vindt men in de eerste rij de getallen 716, 1268 en 2011; dat is de eerste groep waarnemingen, die we hierboven hebben bezien. In de tweede rij staan de getallen 937, 1123 en 1911, dus de tweede groep waarnemingen, die we zojuist bestudeerden. Deze procedure dan is nog 48 maal voortgezet in de successieve rijen onder de tweede. In de laatste kolom vindt men de coëfficiënt *b*, berekend voor elke waarnemingsgroep afzonderlijk. De kolom begint met de hierboven gevonden uitkomsten: 0,8048 en 0,7788, en gaat zo door tot de vijftigste, die 0,8079 bedraagt.

De berekeningen zijn in werkelijkheid nog 450 maal voortgezet (het totale aantal waarnemingsgroepen is dus 500), maar die zijn hier niet gereproduceerd. Het is instructief de resultaten van alle 500 *b*-coëfficiënten grafisch weer te geven; dit is hieronder in de vorm van een frequentieverdeling gedaan. (Voor de in de figuur getekende vloeiende kromme zie blz. 337.)

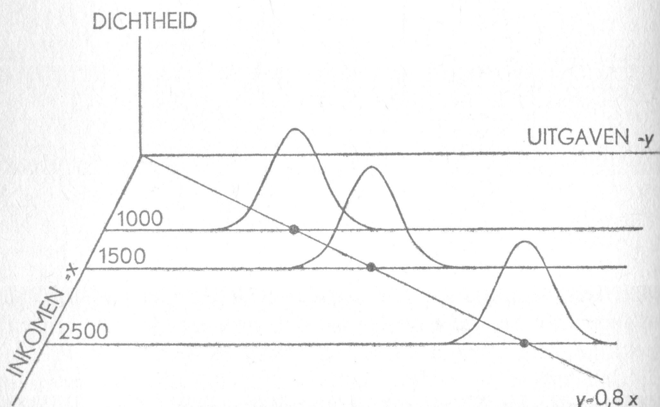


Hoe deze resultaten te interpreteren? De lezer zal al vermoed hebben dat het gepresenteerde het resultaat is van aselect trekken; dit is inderdaad het geval. Wij hebben nl. voor elk van de drie inkomens f 1000, f 1500 en f 2500 precies hetzelfde gedaan als wat wij op blz. 173 alleen voor het inkomen van

f 1000 hebben gedaan. Het komt op het volgende neer. Neem de eerste regel van de tabel op blz. 327. De daar vermelde uitgaafbedragen (716, 1268, 2011) zijn het numerieke resultaat van aselect trekken. Echter niet uit een en dezelfde populatie. Dat zou ook niet voor de hand liggen, want in het eerste geval gaat het om de uitgaven van een gezin met f 1000 inkomen, in het tweede om de uitgaven uit een inkomen van f 1500, en in het derde geval is het inkomen f 2500. Er zijn daarom drie verschillende populaties. Zij zijn alle normaal en hebben alle dezelfde standaarddeviatie (σ) van f 100; maar ze hebben verschillende verwachtingen. Deze zijn als volgt gekozen:

inkomen	verwachting van de uitgaven
f 1000	f 800
f 1500	f 1200
f 2500	f 2000

Dit drietal populaties kan op een eenvoudige manier in een driedimensionaal diagram worden afgebeeld. Daartoe meten we langs twee loodrecht op elkaar staande horizontale assen het inkomen resp. de uitgaven; de verticale as is gereserveerd voor de dichtheid van de drie normale verdelingen. Er zijn slechts drie inkomensniveaux van belang, nl. f 1000, f 1500 en f 2500; deze zijn aangegeven door rechte lijnen in het horizontale vlak evenwijdig aan de uitgaven-as.



Stel nu, dat wij trekken uit de populatie van gezinnen met een inkomen van f 2500. Dan is het resultaat een uitgaafbedrag, dat correspondeert met een zeker punt op de rechte, die bij f 2500 de inkomensas snijdt. Wanneer we te werk gaan volgens de regels van het aselect trekken, dan interpreteren we dit bedrag als een kansvariabele, waarvan de verdeling wordt aangegeven door de normale dichtheid die boven de zojuist besproken rechte is getekend (de 'voorste' normale dichtheidskromme). Voor de twee andere populaties is de afbeelding analoog.

In de figuur zijn voorts drie dikke punten aangegeven. Deze corresponderen met de verwachting van de uitgaven: f 800 bij een inkomen van f 1000 en evenzo f 1200 resp. f 2000 bij inkomens van f 1500 resp. f 2500. Men ziet, dat deze drie punten op een rechte lijn door de oorsprong liggen. Dit is het resultaat van de manier waarop wij de populaties hebben gespecificeerd. Het valt immers gemakkelijk te verifiëren, dat de verwachte uitgaven evenredig zijn genomen met het inkomen; voor elk van de drie inkomensgroepen geldt nl., dat de verwachting van de uitgaven 80% van het inkomen bedraagt. Dat is natuurlijk een bijzondere situatie, maar toch een die van groot belang is voor de regressieanalyse. Beschouw daartoe de regressiecoefficiënt b, die wij i.v.m. hetgeen volgt iets omslachtiger zullen opschrijven:

$$b = \frac{x_1y_1 + x_2y_2 + x_3y_3}{x_1^2 + x_2^2 + x_3^2}$$

$$= \frac{x_1}{x_1^2 + x_2^2 + x_3^2}\, y_1 + \frac{x_2}{x_1^2 + x_2^2 + x_3^2}\, y_2$$

$$+ \frac{x_3}{x_1^2 + x_2^2 + x_3^2}\, y_3 .$$

Nu zijn de inkomensbedragen x_1, x_2, x_3 in onze behandeling eenvoudig vaste getallen, t.w. f 1000, f 1500 en f 2500. Daarentegen zijn de uitgaafbedragen y_1, y_2, y_3 in het geval van aselect trekken kansvariabelen, zolang nog niet in feite getrokken is. Dit betekent, dat b dan ook als een kansvariabele geïnterpreteerd moet worden en wel als een lineaire combinatie van de kansvariabelen y_1, y_2, y_3 met de volgende gewichten:

$$\frac{x_1}{x_1^2 + x_2^2 + x_3^2} = \frac{1000}{9.500.000}$$

$$\frac{x_2}{x_1^2 + x_2^2 + x_3^2} = \frac{1500}{9.500.000}$$

$$\frac{x_3}{x_1^2 + x_2^2 + x_3^2} = \frac{2500}{9.500.000}.$$

Wij weten, dat de verwachting van een lineaire combinatie van kansvariabelen gelijk is aan dezelfde lineaire combinatie van de verwachtingen van die kansvariabelen zelf. Wat die laatste verwachtingen betreft, deze zijn hier genomen op het niveau van 80% van het corresponderende inkomen, dus $0{,}8x_1$ voor y_1, $0{,}8x_2$ voor y_2 en $0{,}8x_3$ voor y_3. Dus is de verwachting van b:

$$\frac{x_1}{x_1^2 + x_2^2 + x_3^2}(0{,}8x_1) + \frac{x_2}{x_1^2 + x_2^2 + x_3^2}(0{,}8x_2) +$$

$$\frac{x_3}{x_1^2 + x_2^2 + x_3^2}(0{,}8x_3) = \frac{0{,}8(x_1^2 + x_2^2 + x_3^2)}{x_1^2 + x_2^2 + x_3^2} = 0{,}8.$$

Maar deze 0,8 is precies dezelfde als de coëfficiënt, die het verband beschrijft tussen de verwachting van de uitgaven en het bijbehorende inkomen!

Nu is de algemene situatie deze, dat de coëfficiënt die het verband tussen inkomen en verwachte uitgaven meet, een onbekend getal is. Deze situatie is geheel analoog aan die van de schatting van het populatiegemiddelde μ. Laten we op overeenkomstige wijze β (spreek uit: bèta) voor dit onbekende getal schrijven, dan is dus βx_1 de verwachting van de uitgaven bij een inkomen van x_1; en analoog βx_2 en βx_3 bij de inkomens x_2 en x_3. De verwachting van de coëfficiënt b wordt nu:

$$\frac{x_1}{x_1^2 + x_2^2 + x_3^2}(\beta x_1) + \frac{x_2}{x_1^2 + x_2^2 + x_3^2}(\beta x_2) +$$

$$\frac{x_3}{x_1^2 + x_2^2 + x_3^2}(\beta x_3) = \beta.$$

Hieruit volgt, dat de verwachting van b gelijk is aan β. Anders gezegd: de coëfficiënt b is een zuivere schatting van de coëfficiënt β, die op zijn beurt de samenhang meet tussen enerzijds het inkomen en anderzijds de mathematische verwachting van de uitgaven. Of algemener: tussen enerzijds de verklarende variabele en anderzijds de verwachting van de te verklaren variabele.

Het is goed even bij dit resultaat stil te staan. Ons uitgangspunt in het begin van dit hoofdstuk was betrekkelijk eenvoudig. Het ging erom een verband van het type $y = bx$ vast te leggen met behulp van een aantal waarnemingsparen voor de variabelen x en y. Dit hebben we gedaan volgens kleinste kwadraten; het resultaat was een eenduidige vastlegging van de coëfficiënt b. Maar later gingen we twijfelen aan die eenduidigheid. Er is geen sprake van twijfel zolang we het waarnemingsmateriaal als gegeven beschouwen; maar dat is juist het belangrijke punt. Het uitgangspunt werd nu veranderd in de zin, dat we de inkomensbedragen (algemener: de waarden aangenomen door de verklarende variabele) nog als gefixeerd bleven beschouwen, de uitgaafbedragen (de waarden van de te verklaren variabele) daarentegen niet. We gingen ons afvragen wat voor uitspraken men over de uitgaven kan doen, gegeven het inkomen; we concludeerden dat dit waarschijnlijkheidsuitspraken moesten zijn. We postuleerden een populatie van uitgaafbedragen voor elk gegeven inkomen, maar deden dit op een zeer speciale manier. De verwachting van het uitgaafbedrag werd nl. evenredig gesteld aan het inkomen; de verwachting werd dus gelijk gesteld aan βx, waarbij x het gegeven inkomen is. De evenredigheidscoëfficiënt β is daarbij een onbekende; onze taak is hem statistisch te schatten. Op dit moment verandert de regressieanalyse van karakter. Oorspronkelijk was zij een methode van aanpassing voor een gegeven waarnemingsmateriaal; zolang dat het geval is spelen waarschijnlijkheidsoverwegingen geen rol en de coëfficiënt b is dan gefixeerd op basis van dat gegeven waarnemingsmateriaal. Maar nu is het β die gefixeerd is, zij het onbekend; het wordt de taak van b ons inlichtingen te verschaffen omtrent die β op precies dezelfde manier als waarop wij tevoren het steekproefgemiddelde m hebben benut om ons in te lichten over het populatiegemiddelde μ. En niet alleen dat de probleemstelling van b en β dezelfde is als die van m en μ, de gevolgde procedure en het resultaat zijn het in feite ook. Immers, wij gingen m interpreteren als een kansvariabele, als het steekproefgemiddelde van een aselecte steekproef, en wij toonden aan dat de verwachting van m dan samenvalt met de te schatten μ. Hetgeen wij kort aangaven door te zeggen dat m zuiver is t.o.v. μ. Evenzo hier: wij zijn b gaan interpreteren als kansvariabele (deze coëfficiënt is immers een lineaire combinatie van uitgaafbedragen die zelf door de nieuwe interpretatie kansvariabelen

zijn geworden); en het bleek dat onder de gemaakte veronderstellingen deze b zuiver is t.o.v. β – dus zuiver t.o.v. de coëfficiënt die de relatie tussen inkomen en verwachte uitgaven pretendeert te meten. Aldus is de regressieanalyse geworden van een methode van aanpassen tot een methode van schatten.

7. STANDAARDFOUTEN

De lezer zal wellicht opmerken, dat de hier beschreven omvorming van de regressieanalyse de methode weliswaar realistischer heeft gemaakt, maar dat anderzijds de variabiliteit van de uitkomsten van steekproef tot steekproef een moeilijk te verteren zaak is. Wat valt er te zeggen als er maar één steekproef is? Zelfs als die getrokken is volgens de regels van de kunst, d.w.z. volgens de methode van aselect trekken, dan nog geeft die ene in feite getrokken steekproef maar één coëfficiënt b één numerieke waarde. Het is natuurlijk prettig te weten – zo zal hij vervolgen –, dat als wij de aselecte procedure ad libitum zouden continueren door keer op keer steekproeven te trekken, het resultaat een onbepaald lange rij van uitkomsten zou zijn waarvan het gemiddelde met β samenvalt. Maar – zo besluit hij – stel eens dat ik inderdaad maar één steekproef heb en dus maar één numerieke b-waarde, wat kan ik dan aan de hand van dát getal zeggen over de samenhang tussen inkomen en verwachte uitgaven?

Ons antwoord zal hem misschien gedeeltelijk teleurstellen. Met zekerheid valt aan de hand van zo'n numerieke b-waarde heel weinig over β te zeggen. Om dit te illustreren is het nuttig terug te keren naar de tabel op blz. 327. Wij weten nu, dat deze tabel geconstrueerd is m.b.v. een β-waarde van 0,8. Nemen we de eerste 5 steekproeven, dan vinden we dat de laagste gerealiseerde b-waarde 0,7788 is; nemen we de eerste 20, dan is de laagste 0,7476; nemen we alle 50 die zijn afgedrukt, dan is de laagste 0,7297. Kennelijk kunnen al dit soort waarden gerealiseerd worden (en nog lagere ook wanneer we maar voort zouden gaan).

Het meest rechtstreekse antwoord moet luiden: de statistische schattingstheorie houdt zich eigenlijk niet bezig met de getalswaarden die steekproeven in concreto opleveren, hoe graag we hier ook zekerheid over zouden willen hebben. Deze

theorie interpreteert de werkelijkheid als de uitkomst van een waarschijnlijkheidsmechanisme en richt zich in de eerste plaats op het mechanisme, niet op de uitkomst. De theorie houdt zich bezig met het ontwerpen van procedures die goede resultaten plegen te leveren in de meerderheid der gevallen; maar meestal is het zo, dat we minder goede resultaten in andere gevallen niet kunnen uitsluiten. Een goede procedure zorgt ervoor, dat de minder goede resultaten zich slechts in uitzonderingsgevallen voordoen, dus in een voldoende gering percentage van alle gevallen. In dit verband is het nuttig nogmaals naar de tabel op blz. 327 terug te gaan. Het is bekend, dat het gemiddelde van alle numerieke *b*'s gelijk zou zijn aan 0,8 als we het aantal onbeperkt lieten oplopen. Hoe staat het nu met de 50 waarden die hier afgedrukt zijn? Welnu, hun gemiddelde is 0,7987; dat ligt dus al dicht bij 0,8. Maar we hebben vermeld, dat er in totaal 500 steekproeven zijn getrokken; dat gaat al een stuk verder in de richting van 'onbeperkt oplopen'. Het ligt daarom voor de hand, dat het gemiddelde van alle 500 coëfficiënten nog dichter bij 0,8 zal liggen. In feite is het 0,8038 en het ligt dus verder van 0,8 dan de 0,7987 van de eerste 50 waarden. Hetgeen onverwacht is, maar bepaald ook weer niet a priori had kunnen worden uitgesloten. Het blijft waar, dat het gemiddelde van de coëfficiënten naar 0,8 convergeert wanneer het aantal onbepaald oploopt; maar het is niet waar, althans niet noodzakelijk waar, dat het verloop van dit convergentieproces zélf regelmatig is. Ook over dit proces kunnen alleen maar waarschijnlijkheidsuitspraken worden gedaan. Wij gebruiken de waarschijnlijkheidsrekening om onzekere situaties te hanteren; maar dit hanteren betekent niet, dat we zekerheid voor onzekerheid in de plaats kunnen stellen. Het betekent alleen maar, dat we de onzekerheid binnen de perken houden door ervoor te zorgen dat de kans op ongelukken voldoende klein is. Het werken met de waarschijnlijkheidsrekening komt neer op een bewust leven met onvermijdelijke onzekerheid, niet op het elimineren van de onzekerheid.

Zojuist spraken wij over de kans op ongelukken. Het is goed hierop nog iets nader in te gaan, omdat het ons in staat stelt uitspraken te doen over de betrouwbaarheid van het met de regressieanalyse behaalde resultaat. We hebben gevonden, dat de regressiecoëfficiënt b (als kansvariabele geïnterpreteerd) zuiver is t.o.v. β, dus dat de afwijkingen $b - \beta$ gemiddeld ge-

nomen nul zijn. Maar dat betekent natuurlijk niet, dat ze individueel nul zijn; en een uitspraak over de orde van grootte van de afwijkingen is daarom kennelijk van belang. Laten we weer terugkeren naar de tabel op blz. 327. De 50 daar gepresenteerde coëfficiënten hebben een laagste waarde van 0,7297 en een hoogste van 0,8760, hetgeen een indruk van de variabiliteit verschaft. Voor hetgeen volgt zal het nuttig zijn met deze maatstaf van de variabiliteit te werken:

$$\sqrt{(1/50)[(b_1 - 0{,}8)^2 + (b_2 - 0{,}8)^2 + \ldots + (b_{50} - 0{,}8)^2]} = 0{,}0329.$$

Hierbij stellen $b_1, b_2, \ldots, b_{50}$ de 50 numerieke uitkomsten voor. Deze worden gemeten in afwijking van 0,8 (de verwachting), gekwadrateerd, opgeteld en gedeeld door het aantal (50). Deze procedure is kennelijk geheel analoog aan die van de variantie (zie blz. 168). Bij de variantie gingen we immers van de waarden aangenomen door de kansvariabele de verwachting aftrekken, de verschillen kwadrateren en vervolgens met de bijbehorende waarschijnlijkheden wegen. In het onderhavige geval zijn de waarschijnlijkheden vervangen door relatieve frequenties; elk van de 50 uitkomsten doet zich eenmaal voor, dus zijn de relatieve frequenties alle 1 / 50. Tenslotte hebben we de vierkantswortel genomen; het resultaat correspondeert dus met de standaarddeviatie.

Nu kunnen we ook de variantie en de standaarddeviatie theoretisch afleiden voor het geval de aselecte procedure niet 50 maar een willekeurig groot aantal *b*-waarden zou hebben opgeleverd. De afleiding heeft plaats in het Aanhangsel van dit hoofdstuk. Het resultaat komt hierop neer, dat de theoretische standaarddeviatie in het onderhavige geval gelijk is aan 0,0324; dus practisch hetzelfde als de hierboven gevonden empirische waarde van 0,0329. Wij zullen ons verder tot de theoretische waarde (de 'standaardfout' van *b*) bepalen, omdat die een grotere principiële betekenis heeft.

Hoe moeten wij dit resultaat plaatsen? Daartoe gaan we weer terug naar het begin. De coëfficiënt *b* is een kansvariabele, heeft dus een verdeling met een verwachting en een standaarddeviatie. Die verwachting is β, want *b* is zuiver; in ons geval is de verwachting 0,8. De standaarddeviatie is de standaardfout 0,0324; dat hebben we zojuist verteld. Neem nu eens aan, dat de verdeling van *b* normaal is. (Aangetoond kan worden dat dit waar is, wanneer de populaties der uitgaafbedragen zelf

normaal verdeeld zijn; de empirische verdeling van blz. 329 lijkt trouwens al op de normale, zoals door de vloeiende kromme is aangegeven.) Wij weten van blz. 178, dat bij een normale verdeling de kans om verder van de verwachting terecht te komen dan tweemaal de standaarddeviatie gelijk is aan 0,0455. In dit geval: de kans dat de aselecte trekkingsprocedure een *b*-waarde oplevert die meer dan tweemaal 0,0324 afwijkt van $\beta = 0{,}8$ is gelijk aan 0,0455. Anders gezegd: de kans is maar 0,0455 dat *b* hoger komt dan 0,865 of lager dan 0,735. Nog anders: de kans is 1 — 0,0455 = 0,9545 dat *b* komt te liggen *tussen* 0,735 en 0,865. Op deze wijze is het mogelijk waarschijnlijkheidsuitspraken te geven omtrent de waarde, die *b* aanneemt, gegeven verwachting (0,8) en standaardfout (0,0324).

Laten we het probleem nu eens omdraaien. In werkelijkheid gaat het er *niet* om een kansuitspraak te geven over de *b* van een aselecte steekproef, gegeven een bekende waarde van β; integendeel, β is in werkelijkheid onbekend en het gaat erom een uitspraak over β te geven met behulp van de aselecte steekproef. Maar ook dit is mogelijk. Daartoe roepen wij weer in herinnering, dat *b* niet meer dan tweemaal de standaardfout ($2 \times 0{,}0324$) van β verschilt behoudens een kans van 0,0455. Dit kan ook aldus worden geformuleerd: als we overwegen een steekproef te trekken volgens de aselecte procedure, dan weten we van tevoren, dat behoudens een kans van 0,0455 de *b*-coëfficiënt van die steekproef minder dan tweemaal de standaarddeviatie van de onbekende β ligt. Voorbeeld: neem de eerste steekproef van de tabel op blz. 327, die dus gekenmerkt is door een *b*-waarde van 0,8048. Doe er tweemaal de standaarddeviatie bij en af, dan vinden wij 0,8696 en 0,7400. De bewering is: de onbekende β ligt tussen deze grenzen – behoudens een kans van 0,0455 dat dit niet waar is! Wij weten dat het wel waar is, want $\beta = 0{,}8$ (maar we zijn wijzer dan men in dit soort situaties pleegt te zijn!). Hadden we dezelfde redenering op steekproef No. 26 toegepast, dan was het misgelopen. We hadden dan geredeneerd, dat β ligt tussen 0,7297 — 0,0648 = 0,6649 en 0,7297 + 0,0648 = 0,7945; wederom behoudens een kans van 0,0455 dat het niet waar is. En het is niet waar zoals we weten; de kans op een foutieve conclusie was minder dan 1 op 20 maar hij heeft zich toch gerealiseerd. Hoe vaak komt dit soort ongelukken voor? Wel, bijna 1 op 20 gemiddeld moet het antwoord luiden.

Op deze manier zien wij, hoe de statisticus het probleem van de onzekerheid tracht te hanteren door de frequentie van foutieve gevolgtrekkingen (i) te bepalen en (ii) klein te houden. Samenvattend komt de procedure op het volgende neer:

Aannemend dat de uitgaven bij elk inkomensbedrag x_1, x_2, . . . normaal verdeeld zijn; dat die populaties een verwachting hebben die evenredig is met het inkomen, dus van het type βx_1, βx_2, . . ., en dat de varianties van die populaties alle dezelfde zijn,

dan is ons *probleem* de schatting van β, dus van de coefficiënt die de samenhang tussen inkomen en verwachte uitgaven meet;

de daartoe gevolgde *procedure* houdt in het aselect trekken van een steekproef, die bij elk inkomensbedrag x_1, x_2, . . . een uitgaafbedrag y_1, y_2, . . . oplevert;

de *eerste berekening* op basis van de uitkomsten van de getrokken steekproef is de bepaling van de regressiecoëfficiënt b (dus 0,8048 wanneer we de eerste steekproef op blz. 327 zouden hebben getrokken);

de *tweede berekening* is de bepaling van de standaardfout (dus 0,0324 in ons geval);

en de *conclusie* luidt, dat 0,8048 een zuivere schatting is van β en dat zijn standaardfout gelijk is aan 0,0324; d.w.z. dat de uitkomst 0,8048 verkregen is door een procedure, waarvoor geldt dat dergelijke uitkomsten soms boven en soms beneden β liggen, maar dat het gemiddelde van die uitkomsten met β samenvalt en hun standaarddeviatie gelijk is aan 0,0324;

en tenslotte, een *implicatie* van deze conclusie (mede gegeven de normaliteit van de verdeling van b) is dat β ligt tussen de grenzen 0,7400 en 0,8696, die verkregen zijn door van de schatting 0,8048 tweemaal de standaardfout 0,0324 af te trekken resp. erbij op te tellen; zulks echter behoudens een kans van 0,0455 dat dit niet waar is, d.w.z. dat β buiten deze grenzen ligt.

Wij besluiten dit hoofdstuk met de volgende opmerkingen:

(1) Wij hebben ons hierboven bepaald tot het geval, waarin men tweemaal de standaardfout bij de schatting optelt en ervan aftrekt. Dat is inderdaad wat het vaakst gebeurt, maar men kan ook bijv. driemaal nemen, hetgeen leidt tot een gereduceerde kans op ongelukken (dus op een interval dat β in feite niet bevat). Aldus speelt men 'safer', maar er moet een prijs betaald

worden: het interval verkregen door driemaal de standaardfout te nemen is natuurlijk wijder, zodat de uitspraak over β minder precies is, meer speelwijdte laat. Evenzo kan men bijv. anderhalf maal de standaardfout nemen, waardoor men meer precisie krijgt maar ten koste van geringere betrouwbaarheid; de kans op ongelukken stijgt dan. Waar het optimum ligt wordt in de eerste plaats bepaald door de vraag, hoe erg het is als β buiten het gekozen interval ligt; d.w.z. door de vraag wat men in concreto met de uitkomst gaat doen.

(2) Gegeven de keuze van tweemaal of driemaal of anderhalf maal de standaardfout, wordt de precisie van de uitkomst bepaald door de grootte van de standaardfout zelf. Nu is op blz. 319 het vermoeden uitgesproken, dat de omvang van het waarnemingsmateriaal een factor is, die de kwaliteit van de uitkomst positief beïnvloedt. Dit is inderdaad het geval. Uit het Aanhangsel blijkt, dat de standaardfout omgekeerd evenredig is met de vierkantswortel van het aantal waarnemingen. Voorts blijkt, dat de standaardfout ook geringer is, naarmate de populaties der uitgaafbedragen (algemener: van de waarden aangenomen door de te verklaren variabele) door minder spreiding zijn gekenmerkt. Dit zijn de populaties waarvoor drie dichtheden getekend zijn op blz. 330.

(3) We hebben ons hier beperkt tot het eenvoudigste geval, nl. dat van een rechte door de oorsprong. Loopt hij niet door de oorsprong, dan is de procedure toch geheel analoog. Bijvoorbeeld, de drie dikke punten van de figuur op blz. 330 liggen dan nog steeds op een rechte (die dan uiteraard niet door de oorsprong gaat). Zijn er meer verklarende variabelen, dan stijgt daarmee het aantal te schatten coëfficiënten. Elk van deze coëfficiënten beschrijft de invloed van een der verklarende variabelen op de verwachting van de te verklaren variabele; elk wordt benaderd door een zuivere schatting en elke schatting heeft zijn standaardfout.

(4) Tenslotte – we moeten niet uit het oog verliezen dat wij onze resultaten hebben afgeleid onder zekere veronderstellingen, die, zoals dat het geval pleegt te zijn, soms wel en soms niet correct zullen zijn. Is de samenhang tussen inkomen en verwachte uitgaven werkelijk lineair? Is de steekproef eerlijk aselect getrokken? Hebben de populaties van de uitgaafbedragen heus alle dezelfde variantie? Zo kan men nog meer vragen stellen en verscheidene daarvan kunnen worden be-

antwoord (meestal met een waarschijnlijkheidsuitspraak!); daarop gaan wij niet in, omdat dit buiten de doelstellingen van dit boek valt. Maar een kleine korrel zout bij de interpretatie van de uitkomsten is wel op zijn plaats en een uitdrukking als 'β ligt tussen 0,7400 en 0,8696 behoudens een kans van 0,0455' kan daarom beter afgerond worden tot 'β ligt tussen 0,74 en 0,87 behoudens een kans van 0,05'.

LITERATUUR

De meeste statistische leerboeken behandelen de regressieanalyse in een betrekkelijk laat stadium. Een voorbeeld is dat van Mood en Graybill [1]; hetzelfde geldt voor de (eenvoudiger) boeken van Moroney [2] en Wijvekate [3]. Het boek van De Wolff [4] is een eenvoudig leerboek voor economen. Johnston [5] richt zich i.h.b. op regressiemethoden; zijn boek is minder elementair dan dat van De Wolff.

[1] Mood, A. M., and F. A. Graybill, *Introduction to the Theory of Statistics*. Second Edition. McGraw-Hill Book Company, Inc., New York. 1963.

[2] Moroney, M. J., *Facts from Figures*. Third Revised Edition. Pelican Book A 236, Penguin Books, Harmondsworth, Middlesex. 1956.

[3] Wijvekate, M. L., *Verklarende statistiek*. Aulaboek Nr. 39, Het Spectrum, Utrecht. 1960.

[4] Wolff, P. de, *Bedrijfsstatistiek*. Vierde herziene en vermeerderde druk. N. Samson N.V., Alphen aan de Rijn. 1955.

[5] Johnston, J., *Econometric Methods*. McGraw-Hill Book Company, Inc., New York. 1963.

AANHANGSEL

Teneinde de standaardfout van een regressiecoëfficiënt af te leiden gaan we er in eerste aanleg van uit, dat er maar twee waarnemingen zijn. Dus y_1 en y_2, beide het resultaat van aselect trekken uit populaties met verwachtingen βx_1 resp. βx_2 en variantie σ^2. De regressiecoëfficiënt is dan

$$b = \frac{x_1}{x_1^2 + x_2^2}\, y_1 + \frac{x_2}{x_1^2 + x_2^2}\, y_2 \,.$$

Gemakkelijk ziet men de juistheid van de volgende vergelijking in:

$$\beta = \frac{x_1}{x_1^2 + x_2^2}\, \beta x_1 + \frac{x_2}{x_1^2 + x_2^2}\, \beta x_2.$$

(Aanwijzing: schrijf de rechterzijde als één breuk met noemer $x_1^2 + x_2^2$.) Door de tweede vergelijking van de eerste af te trekken vinden we dan de volgende uitdrukking voor de afwijking tussen de schatting b en de te schatten coëfficiënt β:

$$b - \beta = \frac{x_1}{x_1^2 + x_2^2}(y_1 - \beta x_1) + \frac{x_2}{x_1^2 + x_2^2}(y_2 - \beta x_2).$$

Het gaat om de variantie van b, die op de bekende manier als volgt wordt gevonden. Schrijf eerst b in afwijking van zijn verwachting (dat is hier al gebeurd, want β is de verwachting van b). Neem dan het kwadraat, dus

$$\text{(1)} \quad (b - \beta)^2 = \frac{x_1^2}{(x_1^2 + x_2^2)^2}(y_1 - \beta x_1)^2 + \frac{x_2^2}{(x_1^2 + x_2^2)^2} \times$$

$$(y_2 - \beta x_2)^2 + \frac{2x_1x_2}{(x_1^2 + x_2^2)^2}(y_1 - \beta x_1)(y_2 - \beta x_2),$$

en neem van dit kwadraat de verwachting, d.w.z. ga de waarden die dit kwadraat aanneemt wegen met de bijbehorende waarschijnlijkheden.

Ons probleem komt dus hierop neer, dat we de verwachting van $(b - \beta)^2$ moeten bepalen, hetgeen blijkens (1) neerkomt op de verwachting van een lineaire combinatie van drie kansvariabelen:

$$(y_1 - \beta x_1)^2, \quad (y_2 - \beta x_2)^2, \quad (y_1 - \beta x_1)(y_2 - \beta x_2)$$

met als respectieve gewichten:

$$\frac{x_1^2}{(x_1^2 + x_2^2)^2}, \quad \frac{x_2^2}{(x_1^2 + x_2^2)^2}, \quad \frac{2x_1x_2}{(x_1^2 + x_2^2)^2}.$$

We weten (zie blz. 171), dat deze verwachting gelijk is aan dezelfde lineaire combinatie van de verwachtingen van de afzonderlijke kansvariabelen. Beschouw dan in de eerste plaats de verwachting van $(y_1 - \beta x_1)^2$. Dat is niets anders dan de variantie van y_1, dus σ^2! Immers, voor de bepaling van die variantie meten we eerst y_1 in afwijking van zijn verwachting, dus we schrijven $y_1 - \beta x_1$ omdat βx_1 de verwachting is van y_1; en dan kwadrateren we $y_1 - \beta x_1$ en wegen met de bijbehorende waarschijnlijkheden, dus we nemen in feite de verwachting van $(y_1 - \beta x_1)^2$. Op geheel overeenkomstige wijze geldt, dat de verwachting van $(y_2 - \beta x_2)^2$ ook gelijk is aan σ^2. We zullen hieronder aantonen, dat de verwachting van het product van $y_1 - \beta x_1$ en $y_2 - \beta x_2$ gelijk is aan nul. Dus hebben we voor de variantie van b:

$$\frac{x_1^2}{(x_1^2 + x_2^2)^2}\sigma^2 + \frac{x_2^2}{(x_1^2 + x_2^2)^2}\sigma^2 + \frac{2x_1x_2}{(x_1^2 + x_2^2)^2}0$$

$$= \frac{\sigma^2(x_1^2 + x_2^2)}{(x_1^2 + x_2^2)^2} = \frac{\sigma^2}{x_1^2 + x_2^2}.$$

Hebben we niet 2 maar n waarnemingen, dan vinden we op overeenkomstige wijze, dat de variantie van de regressiecoëfficiënt gelijk is aan

$$(2) \quad \frac{\sigma^2}{x_1^2 + x_2^2 + \ldots x_n^2}.$$

Voor het in de tekst gebruikte voorbeeld geldt $\sigma = 100$, $x_1 = 1000$, $x_2 = 1500$, $x_3 = 2500$; formule (2) geeft dan $10.000/9.500.000 = 0{,}00105263$ als variantie van b en de standaardfout is de wortel daarvan, dus 0,0324. Merk op, dat de noemer van (2) geschreven kan worden als

$$n \frac{x_1^2 + x_2^2 + \ldots x_n^2}{n}$$

dus als n maal het gemiddelde kwadraat van de x-waarden. Wanneer we ons voorstellen, dat het waarnemingsmateriaal in omvang toeneemt op een dusdanige manier dat dit gemiddelde kwadraat niet verandert, dan stijgt de noemer van (2) evenredig met n. De variantie (2) zelf is dus omgekeerd evenredig met n. En aangezien de standaardfout de wortel daarvan is, is deze omgekeerd evenredig met de wortel van n. Een tweemaal zo kleine standaardfout vereist dus een viermaal zo omvangrijk waarnemingsmateriaal.

Nu dan het bewijs dat de verwachting van het product van $y_1 - \beta x_1$ en $y_2 - \beta x_2$ gelijk nul is. Het gaat om twee kansvariabelen, die beide aselect uit hun respectieve populaties zijn getrokken en dus onafhankelijk verdeeld zijn in de zin van het Aanhangsel van Hoofdstuk 6 (zie blz. 184). Beide kansvariabelen hebben verwachting nul; immers y_1 heeft verwachting βx_1 en door aftrek van βx_1 krijgen we volgens de algebra der verwachtingen (zie blz. 170) een kansvariabele die een verwachting van $\beta x_1 - \beta x_1 = 0$ heeft. Laten we gemakshalve veronderstellen, dat beide variabelen discreet verdeeld zijn, dus met kans p_{ij} de waarden a_i resp. b_j (zegge) aannemen. Onafhankelijkheid wil zeggen $p_{ij} = p_{i.} p_{.j}$ voor alle combinaties van de indices i en j, waarbij $p_{i.}$ en $p_{.j}$ de marginale waarschijnlijkheden zijn. De verwachting van het product is dan (zie ook blz. 183):

$$\sum_i \sum_j p_{ij} a_i b_j = \sum_i p_{i.} a_i \sum_j p_{.j} b_j,$$

d.w.z. de verwachting van de ene kansvariabele maal die van de andere. Maar we zagen zojuist dat beide verwachtingen nul zijn. Dan is ook de verwachting van het product nul.

Nog één punt. Blijkens formule (2) is het voor de berekening van de variantie van de regressiecoëfficiënt noodzakelijk, dat we σ^2 kennen. In de meeste gevallen kennen we σ^2 al evenmin als β. Men pleegt dan met een benaderende procedure te volstaan, die

erop neerkomt dat men σ^2 door de volgende schatting vervangt:

$$\frac{(y_1 - bx_1)^2 + (y_2 - bx_2)^2 + \ldots + (y_n - bx_n)^2}{n - 1},$$

d.w.z. de som van de kwadraten van de afwijkingen rond de regressielijn, gedeeld door $n - 1$. (Zouden we delen door n i.p.v. $n - 1$, dan zouden we de gemiddelde kwadratische afwijking krijgen, maar aan $n - 1$ wordt de voorkeur gegeven omdat dan de schatting zuiver is t.o.v. σ^2.) Toegepast op de eerste steekproef van blz. 327 geeft dit een σ^2-schatting van 5791,54. Substitueren we die in (2), dan leidt dit tot een standaardfout van 0,0247 i.p.v. 0,0324. Dus een flinke afwijking, maar we kunnen nu eenmaal niet verwachten met behulp van slechts drie waarnemingen een betrouwbare schatting van de onbetrouwbaarheid van een regressiecoëfficiënt te krijgen! Men pleegt met grotere aantallen te werken; dan pleegt het resultaat beter te zijn.

13. OM DE GULDEN VAN DE CONSUMENT

1. WAAROM IS HET CONSUMENTENGEDRAG INTERESSANT?

Wij zullen ons in dit hoofdstuk bezig houden met de wijze waarop de consument zijn inkomen besteedt. De variabelen waarom het hier gaat zijn de hoeveelheden die de consument koopt van de diverse artikelen (door economen meestal goederen genoemd), waaruit hij kan kiezen. Deze hoeveelheden zijn voor de consument beslissingsvariabelen. Binnen zekere beperkingen – hem opgelegd door zijn inkomen en door de prijzen die hij voor de diverse goederen moet betalen – kan hij vrijelijk beslissen hoe hij zijn inkomen besteedt.

Men kan in die beslissingen om verschillende redenen geïnteresseerd zijn. Allereerst natuurlijk uit pure nieuwsgierigheid. Maar afgezien daarvan is het bijv. voor de fabrikant van personenauto's van groot belang te kunnen voorspellen hoeveel auto's hij in de komende jaren zal kunnen verkopen. Daarvan hangt immers zijn beslissing af om al dan niet tot uitbreiding over te gaan. En voor de mijndirecties, maar ook voor de leiders van de mijnwerkersbonden, is het nuttig te weten in welk tempo de vraag naar kolen in de toekomst zal afnemen. Ook de overheid komt hier in het spel. In welk tempo zullen er maatregelen nodig zijn om nieuwe werkgelegenheid te scheppen in Zuid-Limburg? (Natuurlijk speelt hier niet alleen de consument een rol, maar ook andere afnemers zoals fabrieken en electrische centrales.)

Een ander voorbeeld (al eerder genoemd in Hoofdstuk 5) is dat van de werkloosheid in de sigarenindustrie, een aantal jaren geleden. Het betrof hier een groot aantal gespecialiseerde, oudere arbeiders die moeilijk elders te plaatsen waren. De regering overwoog toen een accijnsverlaging, waardoor de prijzen van sigaren zouden kunnen worden verlaagd. De bedoeling hiervan was dat de afzet van sigaren en daardoor ook de werkgelegenheid van sigarenmakers zou toenemen. Het effect van deze maatregel was echter afhankelijk van de wijze waarop de consumenten zouden reageren op deze prijsverlaging. Ook hier was er dus aanleiding het gedrag van de consumenten te bestuderen.

Ook voor voorspellingen met behulp van aan- en afvoertabellen (zie Hoofdstuk 3) zijn consumptiestudies van groot belang. Daar worden immers de hoeveelheden die bijv. over drie jaar door de consument verbruikt zullen worden als gegeven beschouwd. De bruikbaarheid van de aan- en afvoeranalyse staat of valt dus met de kwaliteit van de voorspellingen van de te consumeren hoeveelheden.

Het onderwerp van het consumentengedrag is voor ons niet nieuw. We hebben immers in Hoofdstuk 4 bij de behandeling van econometrische macro-modellen met consumptiefuncties gewerkt, die voor het totaal van alle consumenten het totaal van hun verbruik beschrijft afhankelijk van een aantal relevante factoren. Hier echter zullen wij meer gedetailleerd te werk gaan. Weliswaar beschouwen we de consumenten groepsgewijs, dus niet afzonderlijk, maar er zal sprake zijn van detaillering naar afzonderlijke goederen. Dat is trouwens ook noodzakelijk willen we kans maken de hierboven gestelde vragen naar tevredenheid te beantwoorden. Het zal dan blijken dat inkomen en prijzen een belangrijke rol spelen. Een complicerende factor daarbij is dat de gulden geen vaste waarde heeft. Men kon in 1900 ruwweg viermaal zo veel voor een gulden kopen als in 1960. Daarom zal het nodig zijn te werken met begrippen als reëel inkomen en relatieve prijzen. Wij zullen dus eerst die begrippen wat nader bezien.

2. REËEL INKOMEN EN RELATIEVE PRIJZEN

Wanneer iemand een salarisverhoging ontvangt kan hij daar op verschillende manieren op reageren. Om te beginnen kan hij het méér ontvangen bedrag in zijn geheel naar de spaarbank brengen. In dat geval behoeft er voor de rest niet veel te veranderen. Maar in de meeste gezinnen bestaat er wel een lijstje van dingen, die men graag had willen kopen, maar die men zich tevoren niet kon veroorloven. Wij zullen de verleiding weerstaan om hier een opsomming van dergelijke mogelijkheden te geven; de lezer kan al te gemakkelijk voor zichzelf een dergelijke lijst van onvervulde wensen aanleggen. Wij willen slechts wijzen op het aspect van de vervanging dat hierbij een rol speelt. In een gezin waar men graag roomboter eet zal men meestal, als men een pakje roomboter meer koopt, met een

pakje margarine minder volstaan. Evenzo zullen mensen die zich vroeger per trein naar de Middellandse Zee begaven en na de salarisverhoging op vliegen overschakelen, per jaar een geringer aantal treinkilometers 'consumeren'.

Een ander aspect waarop gewezen dient te worden is dat van de geldontwaarding. Wanneer men een loonsverhoging van 5% ontvangt en alle prijzen tegelijkertijd met 5% stijgen, kan men niet meer kopen dan voorheen, en er is dan dus geen reden om veranderingen in het bestedingspatroon aan te brengen. Wanneer echter alle prijzen stijgen, terwijl het inkomen *niet* mee omhoog gaat, dan moet de consument, of hij wil of niet, ergens een veer laten. Hij is gedwongen – we laten de mogelijkheid van minder sparen nu even buiten beschouwing – van bepaalde dingen minder te kopen, of duurdere artikelen te vervangen door goedkopere (bijv. roomboter door margarine).

Wij hebben dus geconstateerd dat niet zozeer het geldbedrag, dat een consument in zijn hand krijgt, van belang is, als wel de hoeveelheden die hij ervoor kan kopen. Dit is voor de economen aanleiding geweest het begrip *reëel inkomen* te introduceren. We zullen dat verder aanduiden met het symbool $\bar{m}$. Het wordt berekend door het *geldinkomen* m (dat is het ontvangen geldbedrag) te delen door een indexcijfer van het prijsniveau, kortweg de prijsindex i_p. Dus

$$\text{reëel inkomen} = \frac{\text{geldinkomen}}{\text{prijsindex}} \text{ of } \bar{m} = \frac{m}{i_p}$$

Stijgt bijv. iemands geldinkomen van f 10.000 per jaar tot f 10.700 (dus met 7%), en stijgt tegelijkertijd het prijsniveau met 4% (zegge van 1 tot 1,04), dan stijgt zijn reële inkomen (uitgedrukt in guldens van het uitgangsjaar) van f 10.000 tot f 10.700 / 1,04 = f 10.288, dus met bijna 3%.

Hoe wordt een dergelijke prijsindex berekend? Wij moeten vooropstellen dat er onder de specialisten op dit terrein uitvoerige discussies over deze kwestie zijn gevoerd, die nog steeds niet tot een eind gekomen zijn. Wij zullen daar echter in dit boek niet op ingaan maar één formule geven, die weliswaar niet aan de allerhoogste maar toch aan redelijke eisen voldoet, en die vanwege zijn eenvoud ook in de praktijk meestal wordt gebruikt. Het betreft de prijsindex van Laspeyres. We bezien hiertoe twee situaties: een uitgangssituatie en een nieuwe situa-

tie die is ontstaan doordat er veranderingen zijn gekomen in de prijzen. Ter onderscheid zullen we de symbolen van de uitgangssituatie voorzien van een ster. Bijvoorbeeld, laat p_1 de prijs van een liter melk zijn in de nieuwe situatie, dan wordt de oude melkprijs aangegeven met p_1^*. Evenzo schrijven we $p_2, p_3, \ldots$ voor de prijzen van de overige goederen in de nieuwe situatie en $p_2^*, p_3^*, \ldots$ voor die in de uitgangssituatie. Bovendien schrijven we $x_1, x_2, x_3, \ldots$ voor de gekochte hoeveelheden in de nieuwe situatie en $x_1^*, x_2^*, x_3^*, \ldots$ voor de 'oude' hoeveelheden. We gaan ons dan afvragen, hoeveel de consument in de nieuwe situatie zou moeten uitgeven als hij precies de oude hoeveelheden $x_1^*, x_2^*, \ldots$ zou willen kopen. Welnu, om evenveel melk te kopen als voorheen (x_1^*) zal hij nu $p_1x_1^*$ op tafel moeten leggen, tegen $p_1^*x_1^*$ voorheen. Zo gaat het ook voor alle andere goederen. Het benodigde bedrag in de nieuwe situatie is dus:

$$p_1x_1^* + p_2x_2^* + p_3x_3^* + \ldots .$$

Dit bedrag gaan we nu vergelijken met het bedrag, dat er in de *uitgangssituatie* nodig was om diezelfde hoeveelheden te kopen:

$$p_1^*x_1^* + p_2^*x_2^* + p_3^*x_3^* + \ldots .$$

Is het nieuwe bedrag hoger dan het oude bedrag (dus de enkel-ster-uitdrukking groter dan de dubbel-ster-uitdrukking), dan zijn gemiddeld genomen de prijzen kennelijk gestegen, in het omgekeerde geval gedaald. Het prijsindexcijfer i_p definiëren we nu als het quotiënt van deze twee bedragen:

$$i_p = \frac{p_1x_1^* + p_2x_2^* + p_3x_3^* + \ldots}{p_1^*x_1^* + p_2^*x_2^* + p_3^*x_3^* + \ldots}.$$

Vinden we bijv. $i_p = 1{,}04$, dan concluderen we tot een stijging van het prijsniveau in de nieuwe situatie van 4% t.o.v. de uitgangssituatie. Vaak wordt het indexcijfer in een andere vorm gepresenteerd, nl. vermeningvuldigd met 100, dus 104 in plaats van 1,04, maar deze verandering is natuurlijk niet essentieel.

Tot nu toe hebben we gekeken naar de beweging van alle prijzen gezamenlijk, zoals die samengevat wordt met behulp van een prijsindex. Maar ook de veranderingen van *afzonderlijke* prijzen zijn van belang in de zin, dat zij de consument in vele gevallen aanleiding geven veranderingen aan te brengen in de hoeveelheden die hij koopt. Dit wordt misschien het duidelijkst geïllustreerd door de verse groenten die in de meeste

Nederlandse gezinnen in de verschillende seizoenen op tafel komen. In de zomer zijn dit onder meer sla, sperciebonen en doperwten, tegen het winterseizoen gaat de huisvrouw over op diverse soorten kool. De zomergroenten zijn dan 'niet meer te betalen'. Meestal nemen we dus een daling van de gekochte hoeveelheid waar, wanneer de prijs stijgt. Deze uitspraak behoeft echter nadere precisering. We moeten namelijk rekening houden met stijgingen van andere prijzen en het inkomen. Wanneer niet alleen de kaasprijs, maar ook alle andere prijzen en het geldinkomen met 5% toenemen, dan is er geen reden om minder kaas te kopen. Het gaat er dus kennelijk om hoe de kaasprijs zich gedraagt t.o.v. alle andere prijzen. Daarom zullen we ook de prijzen moeten delen door een indexcijfer van het prijsniveau. Vaak wordt hiervoor hetzelfde indexcijfer gebruikt als bij de berekening van het reële inkomen. Het quotiënt $\bar{p} = p/i_p$ noemen we dan een *relatieve prijs*. In ons voorbeeld: de kaasprijs in verhouding tot het algemene prijsniveau.

3. HOEVEEL WORDT ER GEKOCHT?

We hebben nu voldoende bouwstenen verzameld om een eenvoudig model te kunnen construeren. We gaan ons afvragen hoe de in Nederland per jaar gekochte hoeveelheid brood afhangt van het reële inkomen en de relatieve prijs van brood. Voor het reële inkomen nemen we in dit geval natuurlijk het reële nationale inkomen per hoofd der bevolking.[1] Maar hoe meten we de hoeveelheid brood? Er zijn vele soorten brood, ieder met een eigen prijs. Natuurlijk zouden we die soorten afzonderlijk kunnen gaan bekijken, maar dan laten de statistieken ons in de steek. We behelpen ons daarom met de volgende procedure. De totale uitgaven aan alle soorten brood bij elkaar genomen zijn bekend. Er bestaat ook een prijsindex van brood, waarin de prijzen van de afzonderlijke soorten zijn opgenomen. Het quotiënt $\bar{x}$ van de totale uitgaven per hoofd en de prijsindex gebruiken we dan als een benadering voor de gekochte hoeveelheid brood.

We beschouwen de gekochte hoeveelheid brood $\bar{x}$, het reële

1. In de practijk neemt men meestal niet het nationale inkomen maar de totale consumptieve bestedingen. Maar wij zullen gemakshalve van inkomen spreken.

inkomen per hoofd $\overline{m}$, en de relatieve prijs van brood $\overline{p}$. Deze laatste is in dit geval verkregen door de prijsindex van brood te delen door de index waarin alle prijzen zijn opgenomen. We gaan nu nog wat verder met deze begrippen manipuleren, door ze om te rekenen tot procentuele veranderingen van jaar tot jaar. Van het reële inkomen $\overline{m}$ in een bepaald jaar trekken we het reële inkomen in het voorafgaande jaar, $\overline{m}_{-1}$, af, delen door dat zelfde reële inkomen in het voorafgaande jaar, en vermenigvuldigen het geheel met 100. Voor het resultaat schrijven we $\overline{\overline{m}}$:

$$\overline{\overline{m}} = 100 \frac{\overline{m} - \overline{m}_{-1}}{\overline{m}_{-1}}.$$

Was bijv. het reële inkomen in het voorafgaande jaar f 1200 en dat in het lopende jaar f 1236, dan vinden we

$$\overline{\overline{m}} = 100 \frac{1236 - 1200}{1200} = 100 \frac{36}{1200} = 100 \times 0{,}03 = 3,$$

dus een stijging van 3 procent. Op analoge wijze bepalen we de procentuele veranderingen $\overline{\overline{x}}$ en $\overline{\overline{p}}$ van resp. de hoeveelheid $\overline{x}$ en de (relatieve) prijs $\overline{p}$:

$$\overline{\overline{x}} = 100 \frac{\overline{x} - \overline{x}_{-1}}{\overline{x}_{-1}}, \quad p = 100 \frac{\overline{p} - \overline{p}_{-1}}{\overline{p}_{-1}}.$$

We gaan nu aannemen dat er tussen deze variabelen ($\overline{\overline{x}}$, $\overline{\overline{m}}$ en $\overline{\overline{p}}$) een lineair verband bestaat. Daarop kan dan de methode der kleinste kwadraten (besproken in het vorige hoofdstuk) worden toegepast, met het volgende resultaat:

$$\overline{\overline{x}} = \underset{(0{,}064)}{0{,}122}\,\overline{\overline{m}} - \underset{(0{,}029)}{0{,}053}\,\overline{\overline{p}}.$$

De tussen haakjes vermelde getallen zijn de standaardfouten, die een maatstaf vormen van de betrouwbaarheid der coëfficiënten.

Hoe interpreteren we dit resultaat? Laten we het eerst houden op de coëfficiënten 0,122 en —0,053, dus doen alsof de onzekerheid, gemeten aan de standaardfouten, geen rol speelt. Dan zegt de vergelijking, dat we de geconstateerde procentuele verandering van het reële inkomen ($\overline{\overline{m}}$) met 0,122 moeten vermenigvuldigen en de procentuele verandering van de relatieve

prijs van brood ($\bar{p}$) met —0,053 om de procentuele verandering van het broodverbruik te vinden. Of ook: ieder procent stijging van het reële inkomen leidt tot een stijging van het broodverbruik met 0,122%. Men pleegt dit uit te drukken door te zeggen, dat de *inkomenselasticiteit* van het broodverbruik gelijk is aan 0,122. Analoog is —0,053 de *prijselasticiteit* van het broodverbruik: ieder procent verhoging van de relatieve prijs van brood leidt, bij gelijkblijvend reëel inkomen, tot een daling van het broodverbruik van 0,053%.[1] De toevoeging 'bij gelijkblijvend reëel inkomen' is niet zonder belang. Want stel dat de broodprijs stijgt, terwijl alle andere prijzen en het *geld*inkomen onveranderd blijven. Dan stijgt het prijsniveau door de stijging van de broodprijs, zodat – gegeven het ongewijzigde geldinkomen – het reële inkomen moet dalen. Het uiteindelijke effect van de gestegen broodprijs is dan tweeërlei: enerzijds het rechtstreekse effect via de relatieve broodprijs (die gestegen is), anderzijds het indirecte effect via het reële inkomen (dat gedaald is).

Dit dan betreft de systematische samenhang tussen enerzijds het broodverbruik en anderzijds het reële inkomen en de relatieve prijs van brood, zoals dat statistisch geschat is. Sinds het vorige hoofdstuk weten we, dat dit soort resultaten aan een tweetal beperkingen onderhevig is. In de eerste plaats zijn er meer factoren die van invloed zijn op het broodverbruik dan inkomen en prijzen alleen, en ook die factoren kunnen fluctueren in het verloop van de tijd. Vandaar dat we spreken van 'systematische samenhang', niet van een volledig verband dat alle bepalende factoren omvat. In de tweede plaats hebben we slechts schattingen van de beïnvloedingscoëfficiënten, die aan steekproeffluctuaties onderhevig zijn. De inkomenselasticiteit van brood wordt geschat op 0,122 en die uitkomst heeft een standaardfout van 0,064. Passen we de regel van 'tweemaal de standaardfout' toe, dus het interval van 0,122 — 2 × 0,064 tot 0,122 + 2 × 0,064, dan luidt de conclusie: de inkomenselasticiteit ligt tussen (ongeveer) 0 en $^1/_4$, behoudens een kans van 1 op 20 dat hij buiten dat interval ligt. U ziet, er is geen sprake van volmaakte precisie.

1. Voor de mathematisch geschoolde lezer: de inkomenselasticiteit is de partiële afgeleide van de logarithme van het verbruik naar de logarithme van het inkomen. Analoog voor de prijselasticiteit.

Wij zullen nu op wat grotere schaal aandacht besteden aan de zojuist ingevoerde inkomenselasticiteiten. Het is dan nuttig met een classificatie naar hun numerieke waarden te beginnen. Neem daartoe allereerst het geval van brood, met een (geschatte) elasticiteitswaarde van 0,122. Stel eens dat het inkomen stijgt met 1%, terwijl alle prijzen onveranderd blijven. Dan wordt dus het inkomen van m tot $1{,}01m$. Het reële inkomen stijgt ook met 1% (de prijzen zijn immers onveranderd) en dus stijgt het broodverbruik met 0,122%. Ook het geldbedrag aan brood uitgegeven stijgt met dit percentage, het wordt dus van (zegge) g tot $1{,}00122g$. Het *aandeel* van het inkomen aan brood besteed was oorspronkelijk g/m, en het wordt nu:

$$\frac{1{,}00122\ g}{1{,}01m} = 0{,}991\frac{g}{m},$$

dus *kleiner* dan het oorspronkelijke aandeel g/m. Deze situatie nu treedt steeds op wanneer de inkomenselasticiteit kleiner is dan 1: iedere procentuele inkomensstijging wordt dan gevolgd door een *geringere* procentuele stijging van het verbruik en dus ook door een geringer aandeel van het goed in het totaal der uitgaven. Omgekeerd, wanneer de inkomenselasticiteit groter dan 1 is, zal het aandeel van het goed bij klimmend inkomen toenemen. Uitgaven voor toeristische doeleinden vormen een voorbeeld. Dergelijke goederen noemt men *luxe goederen*. We spreken van *noodzakelijke goederen* wanneer (zoals bij brood) de inkomenselasticiteit kleiner dan 1 is; die spelen een prominenter rol in de uitgaven naarmate het inkomen geringer is. Een bijzonder geval hiervan zijn de *inferieure goederen*, waarvan niet alleen het aandeel daalt bij klimmend inkomen, maar waarvan bovendien in absolute zin minder verbruikt wordt. Een voorbeeld is margarine; de inkomenselasticiteit is dan negatief. Overigens moeten de termen luxe, noodzakelijk en inferieur niet in waarderende zin geïnterpreteerd worden. Het zijn technische termen, betrekking hebbend op de vraag of de inkomenselasticiteit groter is dan 1, kleiner dan 1 of kleiner dan 0.

We geven nu een overzicht van een aantal door A. P. Barten berekende inkomenselasticiteiten voor een aantal groepen arti-

kelen (gerangschikt naar klimmende waarde van de elasticiteitsschattingen). Ze berusten op statistische gegevens van het C.B.S. voor Nederland uit de jaren 1921-1939 en 1948-1958. Tussen haakjes zijn weer de standaardfouten vermeld.[1]

Brood	0,12	(0,06)
Zuivelproducten	0,56	(0,13)
Tabakswaren	0,62	(0,15)
Banket, chocolade, consumptie-ijs	0,62	(0,15)
Kruidenierswaren	0,67	(0,15)
Verwarming, gas, water, electriciteit	0,69	(0,10)
Aardappelen, groenten, fruit	0,84	(0,17)
Vis en visconserven	0,88	(0,41)
Dranken	0,89	(0,19)
Vlees en vleeswaren	0,92	(0,19)
Schoeisel	1,01	(0,20)
Huishoudelijke artikelen, woninginrichting	1,27	(0,18)
Overige duurzame consumptiegoederen	1,49	(0,23)
Textiel en kleding	1,83	(0,12)

De lijst bevat geen geval van inferieure goederen, want alle uitkomsten zijn positief. (Uiteraard met de kanttekening dat brood daar dichtbij komt, gelet op de standaardfout.) Dit ligt gedeeltelijk aan het feit, dat er nogal sterk gegroepeerd is, waardoor inferieure en niet-inferieure goederen bijeengevoegd zijn met als gevolg dat de groep als geheel niet inferieur is. In een later deel van dit hoofdstuk zullen wij met een fijnere detaillering werken en dan inderdaad enkele negatieve waarden voor de inkomenselasticiteiten vinden.

De lijst laat verder zien, dat waarden boven 1 hoofdzakelijk bij de duurzame goederen zijn geconcentreerd. Dit is wat men ook a priori zou verwachten: stijgt het inkomen, dan pleegt een vrij aanzienlijk deel van die toeneming aan duurzame goederen te worden besteed. De standaardfouten zijn merendeels van de orde van 0,1 à 0,2, hetgeen betekent dat de elasticiteiten gekenmerkt zijn door een onzekerheidsmarge van 0,2 à 0,4. Alleen voor vis en visconserven is de onzekerheid veel groter; wat die artikelengroep betreft helpt het onderzoek ons dus niet veel verder.

1. Ook de elasticiteiten uit de vergelijking voor brood op blz. 349 waren uit dit onderzoek afkomstig.

Wij gaan nu wat meer in detail aandacht besteden aan de prijselasticiteiten. Een nuttig uitgangspunt is de op blz. 349 vermelde vergelijking voor het broodverbruik, die we gemakshalve hier reproduceren:

$$\overline{\overline{x}} = \underset{(0{,}064)}{0{,}122\,\overline{\overline{m}}} - \underset{(0{,}029)}{0{,}053\,\overline{\overline{p}}}.$$

Deze vergelijking bevat slechts één enkele prijs op expliciete wijze, nl. de prijs van brood. Toch spelen de overige prijzen impliciet wel degelijk een rol, en wel op twee manieren. Ten eerste wordt het reële inkomen (de eerste term rechts van het gelijkteken) gevonden door het geldinkomen te delen door een prijsindex, waarin alle prijzen vertegenwoordigd zijn. Ten tweede wordt ook de prijs van brood (de tweede term) op analoge wijze gedeeld door een prijsindex. We kunnen dus zeggen, dat de broodprijs drie keer in de vergelijking voorkomt (één keer rechtstreeks, twee keer indirect via de prijsindex) en alle overige prijzen slechts twee keer.

Het is nuttig om wat nader in te gaan op de betekenis van die indirecte effecten van de prijzen via zo'n prijsindex. Wat de inkomensterm betreft is dit eenvoudig. Een stijging van *alle* prijzen met 1% bij gelijkblijvend geldinkomen is in feite gelijkwaardig met een daling van het inkomen met 1%; we drukten dit al eerder uit door te zeggen dat het reële inkomen dan met 1% daalt. De wijze waarop we de prijzen op het verbruik van invloed doen zijn via de prijsindex van het inkomen is dus niets anders dan een zaak van omrekening. Wat de tweede prijsindex betreft (dus die waardoor de prijs van brood wordt gedeeld), hier ligt de zaak iets subtieler. We moeten ons voorstellen, dat in beginsel alle goederen concurreren om de gulden van de consument. De consument is geïnteresseerd in alle goederen; hoe meer hij ervan kan kopen hoe liever het hem is (tot hij ervan verzadigd is, maar dat punt ligt meestal ver weg). Hij wordt echter gebonden door het feit, dat zijn inkomen nu eenmaal niet hoger is dan het is; en voorts door de prijzen, die voor hem een gegeven zijn. Dit laat zich mathematisch formaliseren aan de hand van een doelstellingsfunctie (meestal nutsfunctie genoemd), die de voorkeuren van de consument

t.o.v. de verschillende goederen weergeeft en die gemaximeerd wordt met inachtneming van de restrictie dat het beschikbare inkomen beperkt is. Het zou te ver voeren op de wiskundige details in te gaan. Wij volstaan met de mededeling, dat juist doordat alle goederen met elkaar om de gulden van de consument concurreren alleen de onderlinge verhoudingen van de prijzen van belang zijn. Vandaar dat de broodprijs gedeeld wordt door een prijsindex.

Maar sommige goederen concurreren meer met elkaar dan andere. Neem sigaretten en shag. Of roomboter en margarine. Het komt ook voor, dat bepaalde goederen eigenlijk nauwelijks met elkaar concurreren maar elkaar eerder aanvullen. Neem lucifers en tabakswaren. Of auto's en benzine. In al die gevallen *blijft* het waar, dat de goederen in het grote geheel met elkaar concurreren, omdat ze immers alle uit hetzelfde inkomen betaald moeten worden; maar *daarnaast* is er sprake van een speciale relatie tussen de twee groepen. Wanneer roomboter duurder wordt en margarine niet, dan mag men verwachten dat er minder roomboter verkocht wordt en tegelijkertijd meer margarine, dus een ontwikkeling van het verbruik in *tegengestelde* richting. Dit is het geval van twee goederen die elkaar's 'substituut' zijn. Wanneer auto's duurder worden heeft dit een negatieve invloed op het aantal gekochte auto's maar ook op het aantal gekochte liters benzine. Dus een effect in *gelijke* richting. We spreken dan van 'complementaire' goederen.

In het hier geciteerde onderzoek gaat het om betrekkelijk omvangrijke groepen van artikelen, zodat dergelijke specifieke effecten dan niet bijzonder duidelijk voor de dag komen. We zullen iets laten zien van een enkel geval, dat in deze richting wijst. Het betreft het verbruik van banket, chocolade en consumptie-ijs, waarvan we de jaarlijkse procentuele verbruiksverandering zullen weergeven met $\bar{\bar{x}}_1$ en de procentuele prijsverandering (gedeeld door de prijsindex) met $\bar{\bar{p}}_1$. De vergelijking luidt als volgt:

$$\underset{}{\bar{\bar{x}}_1} = \underset{(0{,}150)}{0{,}624\bar{\bar{m}}} - \underset{(0{,}108)}{0{,}487\bar{\bar{p}}_1} + \underset{(0{,}052)}{0{,}076\bar{\bar{p}}_2} + \underset{(0{,}079)}{0{,}123\bar{\bar{p}}_3}.$$

De inkomenselasticiteit (geschat) is dus ruim 0,6; de prijselasticiteit 0,5 afgezien van teken. Daarnaast zijn er nog twee andere prijzen: $\bar{\bar{p}}_2$, de procentuele verandering van de prijs van

tabakswaren, en p_3, de procentuele prijsverandering van dranken. De gedachte is, dat de laatste twee goederengroepen substituten zijn voor banket, chocolade en consumptie-ijs. De vergelijking laat inderdaad zien, dat de bijbehorende prijselasticiteiten ('kruiselingse' elasticiteiten genoemd) op positieve waarden geschat worden, nl. 0,076 resp. 0,123. Anderzijds zijn de bijbehorende standaardfouten echter zo groot, dat toepassing van de regel 'tweemaal de standaardfout' ook negatieve waarden toelaat. Men pleegt dit uit te drukken door te zeggen, dat de desbetreffende elasticiteiten 'niet significant verschillend van nul' zijn.

Bij meer gedetailleerde artikelengroepen zal men vaker tot in dit opzicht significante resultaten komen. Gaat het echter om grovere indelingen, zoals hier het geval is, dan kan men meestal met een enkele prijs volstaan, nl. die van de goederengroep in kwestie. Er is dan ook maar een enkele prijselasticiteit per goederengroep, zoals in de vergelijking voor brood. Voor een land als Nederland (en voor andere landen die wat inkomensniveau betreft met het onze vergelijkbaar zijn) geldt, dat een dergelijke prijselasticiteit ruwweg de helft van de bijbehorende inkomenselasticiteit bedraagt.

6. HUISHOUDREKENINGEN

De resultaten die we tot nu toe bekeken hebben waren alle gebaseerd op tijdreeksen. Dat wil zeggen, dat de totale uitgaven in Nederland per goederencategorie gedurende een reeks van jaren werden beschouwd. Er staat ons echter ook een ander type gegevens over het gedrag van consumenten ter beschikking. Deze gegevens worden op de volgende wijze verkregen: aan een groot aantal gezinnen wordt gevraagd hun medewerking te verlenen door al hun uitgaven in een speciaal voor dat doel ontworpen (en dus uniform ingedeeld) huishoudboekje op te schrijven. Natuurlijk is er een bepaald percentage dat zijn medewerking weigert, of de opdracht op inadequate wijze uitvoert. Wat er overblijft pleegt dus niet werkelijk 'aselect' te zijn t.o.v. de bevolking als geheel (of t.o.v. de groepen waarvoor men zich interesseert), zelfs als men gepoogd heeft dit te bereiken bij de selectie vooraf. Een huisvrouw die van zichzelf weet dat zij slordig is doet niet mee; en als zij het is zonder het te weten

komt ze daar spoedig achter en valt zij uit. Zo is het typerend, dat men vaak achteraf heeft kunnen vaststellen, dat de deelnemende gezinnen meer dan normaal plegen uit te geven aan boeken, en minder dan normaal aan alcoholhoudende dranken. Een huishoudrekeningenonderzoek is dus van beperkt nut voor een marktanalyticus als zijn cliënt een bierbrouwerij is; desondanks zijn de verkregen resultaten in de meeste gevallen interessant genoeg om een gedetailleerde analyse te rechtvaardigen. In de meerderheid der gevallen beperkt een dergelijk onderzoek zich tot een periode van 1 of 2 jaar, omdat het moeilijk is voor een langere periode de medewerking van de deelnemende gezinnen te verkrijgen. Dat impliceert dat er in de prijzen zeer weinig variatie valt te bespeuren, althans weinig variatie in het verloop van zo'n beperkte tijd. Er kan wel sprake zijn van *regionale* prijsverschillen, bijv. groenten die in de stad duurder zijn dan op het platteland. Maar dan is het weer bijzonder moeilijk vast te stellen of de kwaliteit wel dezelfde is. Wanneer we dus verschillende gezinnen vergelijken, moeten we niet al te veel illusies hebben over de mogelijkheid hieruit informatie over prijselasticiteiten te halen. Verschillen in inkomens zijn er natuurlijk, en wel in belangrijke mate. Maar daarnaast zijn er nog andere belangrijke factoren, die de verschillen in gezinsuitgaven bepalen: de grootte en de samenstelling van het gezin, en de verschillen in smaak.

Wij zullen nu enige resultaten presenteren van een groot onderzoek naar de koopgewoonten van Britse consumenten, dat werd uitgevoerd in Cambridge onder leiding van Richard Stone. Stone beschikte over een groot aantal huishoudboekjes uit de jaren 1937-1939. Om met de factor 'smaak' te beginnen: het is in een statistisch onderzoek natuurlijk ondoenlijk rekening te houden met alle individuele smaakverschillen. Daarom werden de gezinnen opgedeeld in twee categorieën, arbeiders en middengroepen, om althans iets te doen in de richting van een grotere homogeniteit van de voorkeuren van de consument.

Wat de gezinssamenstelling betreft, is het duidelijk dat bij gelijke inkomens in een groot gezin meer aan voedsel moet worden besteed dan in een klein, zodat er minder voor andere dingen overblijft. Maar een gezin bestaande uit vier volwassenen heeft weer heel andere behoeften dan een gezin dat bestaat uit twee volwassenen en twee kinderen beneden vijf jaar. Daar-

om is met de volgende schaal gewerkt. Een volwassen man werd op 1 gesteld, een volwassen vrouw op 0,9. Voor de kinderen beneden 14 jaar werd er nog weer onderscheid gemaakt tussen arbeidersgezinnen en gezinnen behorende tot de middengroepen. De volgende tabel geeft een overzicht:

	Arbeiders		Middengroepen	
Leeftijdsgroep	mannen	vrouwen	mannen	vrouwen
Beneden 5 jaar	0,52	0,52	0,28	0,28
5 tot 13 jaar	0,52	0,52	0,65	0,65
14 tot 17 jaar	0,98	0,90	0,98	0,90
18 jaar en ouder	1,00	0,90	1,00	0,90

Een arbeidersgezin bestaande uit een ouderpaar, een zoon en een dochter tussen 14 en 17 jaar, en drie kinderen beneden 14 jaar, wordt dus wat consumptie betreft als gelijkwaardig beschouwd met $1{,}00 + 0{,}90 + 0{,}98 + 0{,}90 + 3 \times 0{,}52 = 5{,}34$ volwassen mannen. Natuurlijk is deze schaal bijzonder globaal en valt nauwelijks aan te nemen dat hij correct is voor fruit of melk. Hij is niet meer dan een grof hulpmiddel.

Op basis van de gegevens uit de huishoudboekjes berekende Stone inkomenselasticiteiten. Daarbij moet worden aangetekend dat hij niet van de gekochte hoeveelheden uitging, maar van het uitgegeven bedrag per goederencategorie. Ten eerste is dit sinds het eerste huishoudrekeningenonderzoek (nu meer dan een eeuw geleden) min of meer een traditie, ten tweede zijn in dit geval de berekeningen minder gecompliceerd. Men spreekt in dit geval wel van de inkomenselasticiteiten van de uitgaven, om de tegenstelling aan te geven met de inkomenselasticiteiten van de hoeveelheden die we eerder ontmoet hebben. Ten aanzien van de inkomenselasticiteiten werd door Stone geen onderscheid gemaakt tussen de beide bevolkingsgroepen. Het is echter gebleken dat gezinnen uit de middenklasse van de meeste voedingsmiddelen minder kopen dan wat inkomen en gezinssamenstelling betreft vergelijkbare arbeidersgezinnen. Stone schrijft dit verschijnsel toe aan twee oorzaken. De eerste is dat mensen die physieke arbeid verrichten meer voedsel nodig hebben dan anderen, de tweede dat middengroepgezinnen meer geld uitgeven aan de opleiding van hun kinderen en aan 'cultuur', zodat er minder voor voedsel overblijft. Overigens moet men niet uit het oog verliezen dat de in-

komens in de middengroepen aanzienlijk hoger zijn dan die van de arbeiders.

Nu dan de numerieke resultaten, die alle zijn berekend met behulp van kleinste-kwadratenregressies met als te verklaren variabele de uitgaven per 'volwassen man' en als verklarende variabele het inkomen per volwassen man. Voor het totaal van alle voedingsuitgaven is dan de geschatte inkomenselasticiteit 0,53 met als standaardfout 0,04; dus een hele kleine standaardfout. Inkomenselasticiteiten van het voedingstotaal van $^1/_2$ zijn normaal voor landen met een inkomensniveau als dat van Engeland en Nederland. Voor de gezinnen van de middengroepen ligt het niveau van de totale voedingsuitgaven ongeveer 10% beneden dat van arbeidersgezinnen van overeenkomstige omvang en samenstelling. Voor de afzonderlijke goederengroepen binnen de voedingsgroep blijken de resultaten uit de nu volgende tabel. Ook hier zijn de goederen gerangschikt naar opklimmende waarde van de geschatte inkomenselasticiteit; daarachter staat tussen haken de standaardfout, daarachter of de middengroepen meer dan wel minder verbruiken dan overeenkomstige arbeidersgezinnen.[1]

Taptemelk, karnemelk	—0,59	(0,76)	
Gebakken vis, patates frites	—0,45	(0,51)	
Margarine	—0,16	(0,11)	minder
Bloem	—0,15	(0,11)	
Cacao	—0,10	(0,24)	
Brood	—0,05	(0,04)	minder
Thee	0,04	(0,04)	
Vlees in blik	0,27	(0,12)	
Jam, marmelade	0,30	(0,05)	minder
Rund- en kalfsvlees	0,34	(0,06)	minder
Melk	0,50	(0,18)	
Eieren	0,54	(0,07)	minder
Varkensvlees	0,58	(0,24)	minder
Schapen- en lamsvlees	0,70	(0,12)	
Biscuit	0,71	(0,16)	meer
Vis	0,88	(0,07)	minder
Groenten	0,93	(0,14)	minder
Fruit, noten	1,21	(0,16)	
Koffie	1,42	(0,30)	
Room	1,71	(0,29)	
Maaltijden buitenshuis	2,39	(0,18)	minder

Hier is kennelijk sprake van zes negatieve inkomenselasticiteiten, dus van zes inferieure goederen, waaronder ditmaal brood. Overigens zijn ze geen van alle significant van nul verschillend (toepassing van 'tweemaal de standaardfout' laat in alle zes gevallen positieve waarden toe), maar het zou wel erg toevallig zijn wanneer niet enkele van die zes werkelijk inferieur zouden zijn. Het summum van luxe zijn de maaltijden buitenshuis; dat mag geen verwondering wekken. Dat de middengroepen t.a.v. deze categorie minder uitgeven dan overeenkomstige arbeidersgezinnen moet toegeschreven worden aan het tamelijk veelvuldig gebruik van cantines door de laatste groep. Dat koffie luxe is mag vreemd lijken, maar we zijn dan ook in Engeland! (En de standaardfout is groot.)

7. KWALITEIT EN KWANTITEIT

We hebben reeds vermeld dat men in de analyse van huishoudrekeningen gewoonlijk werkt met uitgaven in plaats van gekochte hoeveelheden. De berekende inkomenselasticiteiten geven dan aan met welk percentage de uitgaven aan een bepaalde goederencategorie stijgen, wanneer het inkomen met 1% stijgt. We moeten er rekening mee houden dat dergelijke inkomenselasticiteiten ons een vertekend beeld kunnen geven, wanneer we werkelijk in hoeveelheden geïnteresseerd zijn. Bij vele goederen is het nl. zo, dat de consument bij een inkomensstijging niet of niet alleen méér gaat kopen, maar ook betere (en dus meestal duurdere) kwaliteiten. Bij gelijke hoeveelheden stijgen dan de uitgaven. Wanneer onze studies zo gedetailleerd zouden zijn dat we bijv. koffiemerken van verschillende kwaliteiten als afzonderlijke goederen zouden beschouwen, zou er geen probleem zijn. Maar meestal is ons cijfermateriaal niet zo uitvoerig en zijn alle koffiemerken gecombineerd. Nog sterker spreekt dit bij een post als rundvlees.

Om dit verschijnsel nader te analyseren zullen we gebruik maken van prijsvariaties. We moeten echter goed in het oog houden wàt voor prijsvariaties. Eén soort houdt zich bezig met de ontwikkeling van de prijzen in het verloop van de tijd;

1. Het meer resp. minder is alleen opgegeven voor die goederen, waarvoor het verschil tussen de twee groepen (hier niet gereproduceerd) groter is dan tweemaal de standaardfout.

dat was onze informatiebron voor het vinden van de prijselasticiteiten. Maar we zagen eerder dat dit soort prijsvariaties bij huishoudrekeningen maar ternauwernood een rol speelt. Dat nu geeft ons de gelegenheid om gebruik te maken van een tweede soort van prijsvariaties, nl. de prijs*verschillen* tussen de verschillende merken, en die te gebruiken als indicatoren voor kwaliteitsverschillen. We mogen immers aannemen, dat als een consument overgaat van een goedkoop koffiemerk naar een duurder merk, hij dit niet doet omdat hij het prettig vindt meer geld aan koffie kwijt te raken, maar omdat hij het duurdere koffiemerk beter vindt. Nu verschaffen huishoudrekeningen maar op zeer beperkte schaal de mogelijkheid na te gaan of en hoe hetzelfde gezin van het ene merk op het andere overgaat; maar wat veel eenvoudiger gedaan kan worden is het nagaan van een samenhang tussen de gemiddelde betaalde prijs per goederengroep (als maatstaf voor de kwaliteit) en het inkomen voor de groep van alle onderzochte gezinnen als geheel.

Laten we dit in iets meer detail bekijken. Neem een gezin, dat een koffiemerk met een prijs van p cent per kg koopt; laat het verbruik q kilo per jaar zijn, zodat het totale bedrag aan koffie per jaar uitgegeven dus pq cent bedraagt. Laat nu het gezin een salarisverhoging van 10% krijgen (of, zo U wilt, beschouw naast dit gezin een tweede gezin met een 10% hoger inkomen). Dan zijn er twee effecten: meer koffie en duurdere koffie. Als de inkomenselasticiteit van de verbruikte *hoeveelheid* $^1/_2$ bedraagt, tijgt die hoeveelheid met $^1/_2 \times 10 = 5\%$. Daarnaast voeren we in de 'inkomenselasticiteit van de kwaliteit', d.w.z. de procentuele verandering van de *betaalde prijs* per procent inkomensstijging. Laat die bijv. $^1/_4$ zijn, dan stijgt de betaalde prijs dus met $^1/_4 \times 10 = 2^1/_2\%$. Hoe werkt dit nu uit voor het in totaal aan koffie uitgegeven bedrag? Heel eenvoudig: eerst was de betaalde prijs p, die is nu $2^1/_2\%$ hoger, dus $1{,}025p$; eerst was de gekochte hoeveelheid q, die is nu 5% hoger, dus $1{,}05q$; het uitgegeven bedrag is dus nu:

$$1{,}025p \times 1{,}05q = 1{,}07625pq,$$

dus ruim $7^1/_2\%$ hoger dan het oorspronkelijk uitgegeven bedrag (pq). Merk op, dat die $7^1/_2\%$ niets anders is dan de $2^1/_2\%$ van de prijsstijging plus de 5% van de hoeveelheidsstijging. Dit nu geldt algemeen: we vinden de inkomenselasticiteit van het uitgegeven bedrag door de inkomenselasticiteiten van de betaalde prijs en van de gekochte hoeveelheid bij elkaar op te

tellen. In ons numerieke voorbeeld: de inkomenselasticiteit van de betaalde prijs is $^1/_4$, die van de gekochte hoeveelheid $^1/_2$, dus die van het uitgegeven bedrag $^1/_4 + ^1/_2 = ^3/_4$; gegeven een inkomensstijging van 10% betekent dit een stijging van het uitgaafbedrag van $^3/_4 \times 10 = 7^1/_2$%.[1]

Het volgende tabelletje, dat betrekking heeft op een huishoudrekeningenonderzoek van Amsterdamse hoofdarbeiders in 1934 en 1935, geeft een aardig beeld.

	Inkomenselasticiteit van de				
Goederencategorie	hoeveelheid		kwaliteit		uitgaven
Vleeswaren	—0,11	(0,03)	0,21	(0,02)	0,11
Thee	0,09	(0,05)	0,04	(0,02)	0,13
Beschuit	0,01	(0,05)	0,21	(0,02)	0,22
Kaas	0,09	(0,07)	0,14	(0,02)	0,23
Chocolade	0,13	(0,13)	0,16	(0,03)	0,29
Vlees	0,07	(0,04)	0,22	(0,03)	0,29
Koffie	0,25	(0,07)	0,10	(0,04)	0,35
Cacao	0,29	(0,15)	0,10	(0,04)	0,40
Koek en gebak	0,31	(0,09)	0,13	(0,03)	0,44

De uitkomsten variëren nogal. Als geheel blijkt echter, dat de inkomenselasticiteiten van de kwaliteit met betrekkelijk kleine standaardfouten waarden aannemen van de orde van 0,1 à 0,2. Voor de negen hier beschouwde goederengroepen zijn de schattingen van de twee soorten inkomenselasticiteiten gemiddeld genomen ongeveer gelijk. De conclusie luidt dus, dat als de consument door inkomensstijging ertoe gebracht wordt meer geld te besteden aan vleeswaren, thee, . . ., koek en gebak, hij ruwweg de helft besteedt aan kwaliteitsverbetering binnen de goederengroep en ruwweg de helft aan meer kilo's per groep.

8. ANDERE FACTOREN WAARDOOR DE CONSUMENT ZICH LAAT LEIDEN

Prijzen en inkomen waren onze belangrijkste hulpmiddelen

1. Dat het niet precies $7^1/_2$ maar 7,625% is lijkt hiermee in tegenspraak, maar is het in feite niet. We hebben nl. wat exactheid aan begrijpelijkheid opgeofferd. Door met infinitesimale i.p.v. 10% veranderingen te werken elimineert men de afwijking.

voor de beschrijving van de ontwikkeling van de consumptieve uitgaven in het verloop van de tijd. Zijn we daarin op adequate wijze geslaagd, dan is het probleem daarmee in feite verlegd. De voorspelling van de consumptieve uitgaven per goederencategorie – een vereiste voor de aan- en afvoeranalyse, zie Hoofdstuk 3 – wordt dan verschoven naar de voorspelling van prijzen en inkomens. Dat kunnen we trachten te doen met behulp van een econometrisch model, zie Hoofdstuk 4, vooropgesteld uiteraard dat het model voldoende gedetailleerd is t.a.v. de prijzen. Dit pleegt bepaald nog niet het geval te zijn, maar de toenemende kwaliteit en kwantiteit van het statistische waarnemingsmateriaal plus de analoge ontwikkeling van de electronische rekenapparatuur geven goede hoop in deze richting.

Toch zijn we bij lange na niet volledig geweest in de opsomming van de factoren, die het consumentengedrag bepalen. Producenten en importeurs besteden jaarlijks miljoenen aan reclame om U en ons ervan te doordringen dat wij hun producten toch vooral moeten kopen. Zij doen dit natuurlijk in de stellige overtuiging dat dit geld goed besteed is. Helaas is er nog maar heel weinig bekend over de kwantitatieve invloed van de reclame op de beslissingen van de consument, hetgeen voor ons een reden is om het probleem aan te stippen, maar er niet verder op in te gaan.

Een ander punt is dat wij er tot nu toe van zijn uitgegaan dat een consument onmiddellijk reageert op prijs- en inkomensveranderingen door zijn uitgaven aan te passen. Er zijn echter verschillende redenen om aan te nemen dat dit niet het geval is. Inkomens uit winsten – of dit nu dividenden van N.V.'s of netto bedrijfswinsten van boeren en kleine zelfstandigen zijn – worden vaak pas achteraf precies becijferd en in het geval van de N.V.'s ook pas achteraf uitgekeerd. De consument die het voornamelijk moet hebben van dergelijke inkomensbronnen heeft dus alle reden om met vertraging te reageren op inkomensveranderingen. Maar ook aan de prijzenkant zijn er dergelijke vertragende factoren. Stel bijv. dat olie voor huisbrand goedkoop wordt in vergelijking met kolen. De consument die nog een goede kolenkachel bezit zal zich echter nog wel twee maal bedenken voor hij tot aanschaf van een oliekachel overgaat. Bij de jonge echtparen die zich nieuw inrichten zal een groter percentage een oliekachel kopen. En bij de

bestaande gezinnen zal dit gebeuren wanneer de kolenkachel ouderdomsgebreken begint te vertonen. De vervanging van kolen door olie vindt dus geleidelijk plaats. Een andere vorm van geleidelijkheid is van psychologische aard. Er zijn consumenten die moeten wennen aan veranderingen van hun reële inkomen, zowel bij stijging als (vooral!) bij daling. Ook prijsveranderingen kunnen vertraagd doorwerken.

Duurzame goederen hebben weer hun eigen problematiek: vervanging en uitbreiding. Als de reparatierekening van Uw auto begint op te lopen denkt U aan een nieuwe. Maar of U die koopt hangt af van wat Uw inkomen is, hoeveel die nieuwe auto kost, wat U zich voorstelt van verdere reparatierekeningen voor de oude auto, wat U allemaal nog verder te betalen heeft. Het ligt weer anders wanneer Mevrouw haar zinnen heeft gezet op een auto voor haarzélf, naast de Uwe. Zo'n kleine rode, die practisch geen benzine kost. En dan komt het moment, waarop zelfs een econoom aanvoelt dat zijn vakkennis niet op alles een antwoord geeft!

LITERATUUR

Een heldere en elementaire uiteenzetting van de theorie en de empirie van het consumentengedrag (in het Nederlands) is geschreven door wijlen de Rotterdamse hoogleraar Koyck [1]. Het meest monumentale werk, maar minder elementair, is dat van Stone [2]; een deel van zijn resultaten is in dit hoofdstuk verwerkt. Voor Barten's onderzoek m.b.v. Nederlandse tijdreeksen zie [3]; een nadere uitwerking is door Theil [4] gegeven. Een overzicht van studies aan de hand van huishoudrekeningen is door Houthakker [5] samengesteld. Deze auteur heeft ook, gelijktijdig met maar onafhankelijk van Theil, de inkomensgevoeligheid van de gekochte kwaliteiten onderzocht [6, 7]. De op blz. 361 vermelde uitkomsten zijn afkomstig van het laatstgenoemde artikel.

[1] Koyck, L. M., 'Consumentengedrag, theorie en empirisch onderzoek'. Deel I van de bundel *Verbruik en sparen in theorie en practijk*, samengesteld ter gelegenheid van het vijfenzeventig jarig bestaan van de Rijkspostspaarbank en uitgegeven door H.D. Tjeenk Willink en Zoon, Haarlem, voor het Staatsbedrijf der P.T.T. 1956.

[2] Stone, R., *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom, 1920-1938*. Volume I. University Press, Cambridge (Engl.). 1954.

[3] Barten, A. P., Consumer Demand Functions under

Conditions of Almost Additive Preferences'. *Econometrica*, Vol. 32 (1964), te verschijnen.

[4] Theil, H., 'The Information Approach to Demand Analysis'. Rapport 6310 van het Econometrisch Instituut van de Nederlandsche Economische Hoogeschool (1963).

[5] Houthakker, H. S., 'An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel's Law". *Econometrica*, Vol. 25 (1957), pp. 532-551.

[6] Houthakker, H. S., 'Compensated Changes in Quantities and Qualities Consumed'. *Review of Economic Studies*, Vol. 19 (1951-52), pp. 155-164.

[7] Theil, H., 'Qualities, Prices, and Budget Enquiries'. *Review of Economic Studies*, Vol. 19 (1951-52), pp. 129-147.

EPILOOG

Het begin van dit boek en het einde betreffen beide de persoon om wie het uiteindelijk gaat in de economische problematiek: de consument. Wij begonnen met een huisvrouw, die zich tot taak stelt haar uitgaven aan voedingsmiddelen zo gering mogelijk te doen zijn, echter met inachtneming van de restricties die de goede gezondheid van haar gezin haar oplegt. En wij eindigden in Hoofdstuk 13 met een veel omvangrijker probleem: hoe besteedt de consument zijn inkomen aan de vele artikelen die te koop zijn, hoe verandert zijn bestedingspatroon wanneer zijn inkomen stijgt of wanneer hij hogere prijzen moet betalen, wat zijn de andere belangrijke factoren die het consumptieve gedrag bepalen? Maar ook tussen het begin en het einde van het boek speelde de consument zijn rol: in het hoofdstuk over waarschijnlijkheidsrekening, toen we de variabiliteit van de gezinsuitgaven bekeken; in het hoofdstuk over de statistische specificatie van economische relaties, toen het ging over die variabiliteit bij verschillende inkomensniveaux; en in het hoofdstuk over econometrische macro-modellen, waar een centrale rol werd gespeeld door de factoren die de totale consumptie van een land als geheel bepalen.

Overigens: een consument consumeert niet alleen, hij doet tal van andere dingen; wat we daarvan hebben laten zien was soms tijdrovend, soms luchthartig. Tijdrovend was het wachten bij de kapper en in het postkantoor; wij hebben gepoogd zijn lot in dit opzicht te verbeteren door voorrangsregelingen e.d. voor te stellen. Luchthartig was de quiz, het wedden op Duindigt, de voetbalpool, het spel van kruis of munt, van morra uit Genua, van boter-kaas-en-eieren, en (tot op zekere hoogte) schaken. We introduceerden deze onderwerpen niet onder het motto l'art pour l'art maar om de toepasbaarheid van hun oplossingsprocedures op soortgelijke situaties bij bedrijven en andere organisaties. Wat de luchthartige speltheorie betreft, we lieten eerst twee luchtvaartmaatschappijen met elkaar concurreren en later zagen wij onze landbouwers geconfronteerd met hun varkensfokproblemen. Het schaken bracht ons op het strategiebegrip; dat bracht ons op een rationele manier om onze overtollige geldmiddelen te beleggen. Het strategiebegrip leidde ook tot een rationele aanpak voor de aan-

passing van productie, voorraden en personeelsbezetting van een verffabriek in het verloop van de tijd. Het is inderdaad waar, dat het grootste deel van dit boek gewijd is aan bedrijfsproblemen. Er was een fabrikant die zich afvroeg: hoe maak ik de hoogste winst op mijn radio's? Een tweede vroeg zich af: hoe verdeel ik mijn productie over de successieve kwartalen zodanig dat voorraadkosten plus instelkosten zo gering mogelijk zijn? Een derde: in welke series dien ik te produceren om zo min mogelijk kosten te maken? Een vierde: hoeveel monteurs moet ik in dienst nemen om mijn machines te laten repareren, gegeven dat het geld kost om een defecte machine op reparatie te laten wachten en dat het óók geld kost meer monteurs in dienst te hebben dan er defecte machines zijn? Dat is een wachttijdprobleem, analoog aan de kapper en het postkantoor. En evenzo is de quiz analoog aan tal van transportproblemen (zoals die van onze brandweerauto's), omdat het wiskundig bezien in beide gevallen gaat om een lineair programmeringsvraagstuk.

Sommige van de behandelde problemen zijn macro-economisch van aard, hebben dus betrekking op grote groepen van bedrijven of gezinnen, sommige zijn van het micro-economische type. Dat geldt in de consumptieve sfeer: de afzonderlijke huisvrouw heeft micro-economische problemen (wat nog niet betekent dat zij ze als kleine problemen zal beschouwen), maar de factoren die de totale consumptie van een land bepalen vormen een macro-economisch probleem. Hetzelfde onderscheid is ook van toepassing in de productieve sfeer. De aan- en afvoeranalyse is macro-economisch van aard: het gaat daar om hetgeen geleverd wordt door groepen van bedrijven in sectoren gerangschikt, dus bijv. de landbouw, de textielnijverheid, de chemische nijverheid, enz. De kritieke-padmethode is daarentegen weer volstrekt micro: het gaat om een bepaald project, waarvan de afzonderlijke deeltaken gecoördineerd worden door een enkel of evt. een paar bedrijven. Nog meer macro dan de aan- en afvoeranalyse is bijv. een investeringsvergelijking van een econometrisch model, die de factoren tracht aan te geven waardoor het totaal van alle investeringen van alle bedrijven in een zeker land wordt bepaald. Micro-economisch van opzet is de investeringsenquête van het C.B.S., omdat de basisgegevens door afzonderlijke bedrijven worden verstrekt. De enquête in gepubliceerde vorm is echter macro (het betreft

groepen van bedrijven), en levert dus macro-economische voorspellingen van de investeringen in het komende jaar. Ook de Centrale Economische Plannen zijn macro-economische voorspellingen. We hebben gezien dat ze de neiging hebben de veranderingen in de nabije toekomst te onderschatten, dus te veel gewicht toe te kennen aan het heden en het recente verleden, net zoals sportjournalisten de neiging hebben om bij hun tips voor de voetbaluitslagen op de komende zondag te veel de nadruk te leggen op de prestaties van de elftallen in de afgelopen weken.

Die onvolmaakte voorspellingen brachten ons ertoe aan het probleem van de onzekerheid expliciete aandacht te besteden. Dit leidde tot het hoofdstuk over waarschijnlijkheidsrekening, dat blijkens het overzicht op blz. 12 een soort scharnier vormt in het boek als geheel: bijna alle hoofdstukken daarna hebben dit onderwerp als wiskundige basis. Zoals voor vrijwel alle toepassingen van vrijwel alle onderwerpen van de wiskunde geldt, is een zekere mate van vereenvoudiging van de werkelijkheid nodig opdat de waarschijnlijkheidsrekening op concrete situaties kan worden toegepast. Maar soms dreigt dit te ver te gaan; de vervormde werkelijkheid houdt dan op een voldoende benadering van de werkelijke werkelijkheid te zijn om nog voldoende belangwekkend te zijn. Dan kan men gaan simuleren, d.w.z. een waarschijnlijkheidstheoretische afbeelding maken op meer empirische basis. We kunnen zeggen dat simulatie een soort eerherstel is van het aloude getallenvoorbeeld, dat meer illustratie dan bewijs pleegt te zijn, maar dan veel intelligenter en op veel grotere schaal. We kunnen ook zeggen, dat simulatie een soort substitutie van hooggekwalificeerde arbeid door de diensten van een hooggekwalificeerde machine is: in plaats van het denkwerk benodigd voor de benadering van de werkelijkheid door een wiskundig model laat men een electronische rekenmachine optreden. (Ter geruststelling zij verzekerd, dat er genoeg denkwerk overblijft.) Overigens is er simulatie in uitgebreider zin, waarbij niet of althans niet noodzakelijk van de waarschijnlijkheidsrekening wordt gebruik gemaakt: het toetsen van schepen en dijken in waterloopkundige laboratoria, het houden van beleidsspelen waarin de faits et gestes van enkele bedrijven worden nagebootst, enz.

Wat is er nu voor gemeenschappelijks aan deze variëteit van onderwerpen? In eerste aanleg lijkt dat ver te zoeken, maar bij

iets dieper graven valt gemakkelijk in te zien dat die onderwerpen op twee manieren aan elkaar gebonden zijn. In de eerste plaats de methode van onderzoek: die is duidelijk wiskundig georiënteerd. In de tweede plaats het terrein van onderzoek: onze voorbeelden lagen op economisch terrein, hetzij micro hetzij macro. Toch moeten wij wat dit laatste betreft een beperkende kanttekening maken. Het moge zo zijn dat de overgrote meerderheid van de toepassingen op economisch terrein ligt, het is niet gezegd dat dit zo zal blijven. Integendeel, de ontwikkeling van wat in de Angelsaksische literatuur 'behavioural science' heet suggereert, dat de toepasssing van wiskundige technieken op andere wetenschappen die het menselijk gedrag tot object hebben eveneens vordering zal maken. Er zijn zelfs practische toepassingen, ook al wordt dit door de betrokkenen niet als zodanig gevoeld. Neem bijv. een voorwaardelijke veroordeling. Dat is een strategie! Een strategie geeft immers aan wat in de toekomst te doen staat afhankelijk van de informatie die in de tussentijd verkregen zal zijn. Welnu, een voorwaardelijke veroordeling is van het type 'Gij zult twee jaar zitten als ge U binnen vijf jaar aan herhaling van Uw delict schuldig maakt – zo niet, dan zult ge een vrij man blijven' en dus is een voorwaardelijke veroordeling inderdaad een strategie.

Door aan de realisatie van de voorwaarde een waarschijnlijkheidsuitspraak te verbinden zijn we in staat een voorspelling te geven van feitelijk toekomstig gedrag. Laat bijv. de kans op de zojuist genoemde herhaling 0,3 bedragen, dan is dat ook de kans dat die twee jaar zullen moeten worden uitgezeten. Het was op deze basis, dat wij in Hoofdstuk 11 in staat waren de verffabrikant zijn eigen toekomstige handelingen (inzake productieniveau en de omvang van zijn personeel) te laten voorspellen. Dat waren zuivere voorspellingen, d.w.z. voorspellingen zodanig dat de afwijkingen een mathematische verwachting gelijk nul hebben. In Hoofdstuk 5 verlangden wij van 'wetenschappelijke' voorspellingen een kansuitspraak over hun relatie tot het bijbehorende voorspelde verschijnsel. Dat viel toen niet nader te preciseren, omdat de waarschijnlijkheidsrekening nog niet aan de orde was geweest. Die restrictie is nu vervallen en een zuivere voorspelling vormt een voorbeeld van wat met een kansuitspraak wordt bedoeld. Niet meer dan een voorbeeld, want er zijn tal van andere kansuitspraken denkbaar. Bijv. een 'conservatieve' voorspelling, die – naar de mening van de voor-

speller – met kans $^3/_4$ te laag zal blijken en met kans $^1/_4$ te hoog.

Tot slot nog dit: valt er op basis van de achter ons liggende hoofdstukken een algemene uitspraak te doen over de te volgen procedure bij rationeel handelen aan de hand van wiskundige technieken? Wanneer de vraag zo algemeen wordt gesteld is het antwoord noodzakelijkerwijs wat vaag, hetgeen ietwat frustrerend is voor een vak dat zich juist ten doel stelt concrete antwoorden te geven. Maar toch is het wel mogelijk iets verstandigs op deze vraag ten antwoord te geven. Om te beginnen dan is er steeds een probleem: schepen blijven te lang liggen voordat de laad- en losfaciliteiten beschikbaar komen, de voorraadkosten van een bepaald product lijken excessief hoog, of wat dan ook. Men gaat dan na door welke wiskundige verbanden het probleem gekenmerkt is (of de tussenaankomsttijden van de schepen exponentieel verdeeld zijn, hoe de voorraadkosten met de omvang van de voorraad variëren, enz.). Vervolgens bezint men zich op de vraag welke maatregelen zo al genomen zouden kunnen worden om iets aan het probleem te doen (extra laad- en losfaciliteiten, een prioriteitsregeling, een extra pakhuis huren of laten bouwen, de productie meer met de vraag laten op en neer bewegen, enz.). Die maatregelen zullen zekere effecten hebben, soms ook neveneffecten; ook dat dient bij voorkeur wiskundig beschreven te worden en bovendien dient men rekening te houden met de kosten van de maatregel als zodanig. Daarna komt het probleem van het afwegen. Moeten er op basis van de berekende effecten zekere maatregelen genomen worden en, zo ja, welke en in welke mate? Dit vereist een duidelijk gedefinieerd criterium (winst, kosten, wachttijd, marktaandeel, enz.). Aan dit criterium zal men zo goed mogelijk wensen te voldoen, hetgeen in het algemeen zal leiden tot een bepaalde maatregel die dan de beste blijkt te zijn. Soms komt het voor dat 'lijkt' beter is dan 'blijkt', omdat die beste maatregel bij verder doorrekenen onverwachte en onaangename nevengevolgen blijkt te hebben. Dat betekent niets anders dan dat het probleem niet goed is gesteld en men dient dan de zaak opnieuw op te zetten, maar nu beter. Hoe dan ook, na een of meer ronden komt men dan tot de maatregel die we de beste zullen noemen.

Daarmee is de kous nog niet af. Wij willen het hier niet hebben over de vraag hoe de aldus berekende maatregel in feite wordt doorgevoerd. Dat is vaak een kwestie van overreding,

omdat degene die dit soort wiskundig werk uitvoert in de meeste gevallen niet dezelfde is als degene die de maatregel heeft te nemen. Laten we echter aannemen, dat de als de 'de beste' aangemerkte maatregel wordt uitgevoerd. We dienen dan te bedenken, dat 'best' gebaseerd is op berekeningen, dus in feite op voorspellingen, en dat de toets van de realiteit nu pas komt. En het kan wel eens tegenvallen, hetzij omdat de wiskundige verbanden het mechanisme van de werkelijkheid toch niet helemaal adequaat beschrijven, hetzij omdat het gehanteerde criterium onjuist of onvolledig is. Men begint dan opnieuw en tegelijk doemen andere problemen aan de horizon op – dat is wat de beoefenaars van het vak tot druk bezette mensen maakt.

REGISTER

(De term *inkomenselasticiteit* is steeds afgekort tot *ink. el.*)

Aan- en afvoeranalyse,
– algemeen, 77-100
– voorspellingskwaliteit, 147-154
aan- en afvoertabel,
– algemeen, 79, 80
– Nederlandse, 81-84
aantallen, wet van de grote, 257
aardappelen, ink. el. 352
afstandenschema, 57, 50-66
aggregaten, 101
aggregeren, 78
aselect trekken, 323
assen(kruis), 20
autofabrikant, 188-203

Banket,
– ink. el., 352
– vraagfunctie, 354, 355
BARTEN, A. P., 306, 351, 363
bedieningsinstallatie, 229, 233
bedieningstijd, 233, 234
bedrijfseconometrie, 10
bedrijfsspelen, *zie* beleidsspelen
beginoplossing, 32, 33, 58-66
beheerste variabele, 16, 17
belastingverlaging, gevolgen van, 103, 119
beleidsspelen,
– algemeen, 258, 259
– bromfietsenfabriek, 259-266
beschuit, ink. el., 361
besliskunde, 10
beslissingsregel (*zie ook* strategie),
– algemeen, 186, 187
– lineaire, 290, 292, 294-305
beslissingsspelen, 259
beslissingsvariabele, 17, 18, 27-31, 37, 50, 344
bestedingen, consumptieve, 94
(*zie ook* consument)
betaalrooster, 193, 205
betalingsbalans, 101, 103
– effect van overheidsuitgaven op, 122, 123
bezettingsgraad, 234-241
biscuit, ink. el., 358
bloem, ink. el., 358
BOAS, J., 238
BOEHM, G. A. W., 76
BOGAARD, P. J. M. VAN DEN, 306
BOOT, J. C. G., 124, 125
'bottlenecks', 69, 282
brandweer, 49-67, 68
bromfietsen, 259-266
brood,
– ink. el., 352, 358
– vraagfunctie, 353
BURG, A. R. VAN DEN, 305
business games, 259

Cacao, ink. el., 358, 361
capaciteit, 18, 99
– benutten van, 33, 34
– waarde van eenheid, 34
capaciteitsgrens, 18
capaciteitsrestrictie, 19-23, 29
C.B.S., 81, 94-96, 100, 131-139, 154, 155
CHENERY, H. B., 100
chocolade,
– ink. el., 352, 361
– vraagfunctie, 354, 355
CHRIST, C. F., 154, 155
CLARK, P. G., 100
COBHAM, A., 245
coëfficiënt, technische, 88-98
COHEN, K. J., 267
combinatie, lineaire, 171, 325
concentraties, gevolgen van –

voor aan- en afvoertabel, 84
conjunctuurtest, 154, 155
constante, 10
consument, consumeren, consumptie, consumptieve bestedingen:
– in aan- en afvoeranalyse, 78-88, 94, 97
– van afzonderlijke goederen, 344-363
– in macromodel, 101, 105-109
consumentengedrag, 344-363
consumptiefunctie, 105, 108
consumptieprijspeil, 123
consumptie-ijs,
– ink. el., 352
– vraagfunctie, 354, 355
C.P.M., 69, 70, 76
Critical Path Method, 69

DANTZIG, G. B., 47, 48
DANTZIG, D. VAN, 154, 155
definitievergelijking, 111
degeneratie, 48, 55n.
diagonaal element, 84
dichtheid(sfunctie), 175
dieetprobleem, 14-17, 30, 47
diensten(sector), 78-100
doelstellingsfunctie,
(*zie ook* minimering *en* maximering)
– algemeen, 15, 24, 32, 37, 41, 50, 58-64, 68
– van brandweerhoofdman, 50
– van consument, 353
– op Duindigt, 37, 38
– van fabrikant, 41
– kwadratische, 48
– lineaire, 17, 19, 68
– van radiofabrikant, 18
dranken, 82-84, 95-97
Duindigt, 36-39, 47, 159
duurzame consumptiegoederen, overige, ink. el., 352

Econometric Society, 10
Econometrisch Instituut, 72, 73
econometrisch model, 104
eieren, ink. el., 358
electriciteit, ink. el., 352
electronische rekenmachine, 9, 71, 94, 251, 261, 362
endogeen, 106
ERLANG, A. K., 228
exogeen, 106, 112
extrapolatie, lineaire, 130

FELLER, W., 160, 164, 180, 245
Financiën, Minister van, 103, 117, 118, 126
FORD, L. R., 75, 76
frequentie, relatieve, 160-163, 179, 230, 231
FREUDENTHAL, H., 180, 193
fruit, ink. el., 352, 358
FULKERSON, D. R., 75, 76
functie,
– kwadratische, 15n.
– lineaire, 15n.

GALTON, F., 318
Gamma E.T., 251
gas, ink. el., 352
GASS, S. I., 48, 75
GAUSS, C. F., 315
gebak, ink. el., 361
gebeurtenis, 161, 162
– onmogelijke, 161, 162
– zekere, 161, 162
gebied, toelaatbaar, 24, 27, 29
gedragsvergelijking, 111
geldontwaarding, 345, 346
gemiddelde (gewogen), 167
(*zie ook* verwachting)
gevoeligheidsanalyse, 274-277
gezinsinkomen, *zie* inkomen (van consument)
GOLDBERGER, A. S., 124
GOUDRIAAN, J., 305
grafische voorstelling, 20

– van restricties, 21-23
– van lineaire vergelijkingen, 22, 23
GRAYBILL, F. A., 340
GREENE, J. R., 267
grens (van toelaatbaar gebied), 27, 29
groenten, ink. el., 352, 358
GUETZKOW, H., 267

HADLEY, G., 305, 306
HARRIS, F., 305
HEES, R. N. VAN, 305
HITCHCOCK, F. L., 75
hoekpunt (van toelaatbaar gebied), 27, 29, 31
HOFFMAN, A. J., 48
HOLT, C. C., 278, 300, 305, 306
HOUTHAKKER, H. S., 363, 364
hout(industrie), 82-84, 95-97
huishoudelijke artikelen, ink. el., 352
huishoudrekeningen, 355-359
HULSHOF, A. H., 267

Import, *zie* invoer
industrie (sector), 78-100
inferieur goed, 351
inkomen,
– van consument, 172-175, 319-322, 326-331, 345-352, 358-361
– nationaal, 101, 105, 108
– sector-, 80, 81, 99
inkomenselasticiteit,
– inleiding, 350
– van de hoeveelheid, 350-352
– van de kwaliteit, 360, 361
– van de uitgaven, 358-361
input-output table, 79
zie ook aan- en afvoertabel
instelkosten, 41-46
zie ook omschakelingskosten
intervalvoorspellingen, 128
investeringen
– per bedrijfstak, 132-139
– totale, 101, 105-110
investeringsenquête, 131-139
investeringsgoederen, geïnstalleerd vs. besteld, 136, 137
investeringsvergelijking, 109, 110
invoer, 81, 121, 122
invoerprijspeil, effect op consumptieprijspeil, 123

JACOBS, W. W., 48
jam, ink. el., 358
JOCHEMS, D. B., 154, 155
JOHNSTON, J., 340

Kaas, ink. el., 361
kalfsvlees, ink. el., 358
kans, *zie* waarschijnlijkheid
kansvariabele(n), 165
– onafhankelijkheid van, 184
KANTOROVICH, L., 47, 48
kapitaalgoederenvoorraad, 110
karnemelk, ink. el., 358
KIBBEE, J. M., 267
klant, bij wachttijdprobleem, 229
kleding, ink. el., 352
kleding(industrie), 82-84, 95-97
KLEIN, L. R., 108, 110, 124
kleinste kwadraten, methode der, 315-318
knelpunten, 69-72, 98-100, 282
knooppunten, 70
koek, ink. el., 361
Koekoek, *zie* Nachtegaal
KOERTS, J., 245
koffie, ink. el., 358, 361
KOOPMANS, T. C., 48, 75
kosten, *zie onder* instelkosten, minimering, omschakelingskosten, productiekosten, voorraadkosten
kostenfunctie, kwadratische, 290
KOYCK, L. M., 363
KRAFT, C. J., 267

kritieke pad, 69-76
kruidenierswaren, ink. el., 352
kruis of munt, 206, 214-219
KUHN, H. W., 76
kwadratische functie, 15
kwantitatief-economische studierichting, 11

Lamsvlees, ink. el., 358
landbouw(sector), 78-100
LASPEYRES, E., 346
leegloop, 72, 227, 228, 281, 282
LEONTIEF, W. W., 78, 100
leveringen,
– onderlinge, 79, 85, 98
– binnen sector, 84
limietstelling, centrale, 179
lineair programmeren, 14, 19, 47, 48, 75, 76, 223n.
lineaire beslissingsregels, 290, 292, 294-305
lineaire functie, 15n.
lineaire vergelijkingen, 10, 24, 92, 93
LIPS, J., 140, 154, 155
LIPS, L., 76
looninkomen, 108, 109
loonkosten, normale, 278, 279
loonsomvergelijking, 110
loonvoet, effect op consumptieprijspeil, 123
lopende band, 18, 19
LUCE, R. D., 203, 226
luchtvaartmaatschappijen, 204-206, 213
luxe goed, 351

Maaltijden buitenshuis, ink. el., 358
machinehal, 242-245
macro-economie, 101
macro-modellen, 101, 306
management games, 259
margarine, ink. el., 358
margarinefabriek, 248-254
marktanalyse, 86
marktanalyticus, 356
marmelade, ink. el., 358
massafunctie, 165
maximering,
– algemeen, 17, 18
– van laagste verwachte winst, 217-224
– van laagste winst, 36-39
– van nutsfunctie, 353, 354
– van prestatie van quiz-team, 67-69
– van verwachte winst, 194-202
– van winst, 18-39
maximumprobleem, *zie* maximering
MCKINSEY, J. C. C., 226
MEERENDONK, H. W. VAN DEN, 305
melk, ink. el., 358
menu, goedkoopste, 14
meubelindustrie, 82-84, 95-97
minimax,
– criterium, 38, 212
– stelling, 222, 219-224
minimering,
– algemeen, 17, 18
– van af te leggen afstand, 49-67
– van grootste verwachte verlies, 221
– van kosten van dieet, 14-17
– van kwadratensom van afwijkingen, 315-318
– van leegloop- en wachtkosten, 248, 254
– van productiekosten, 39-47
– van totale kosten, 285-290
– van transportkosten, 52
– van verwachting totale kosten, 288
minimumprobleem, *zie* minimering
model,
– dynamisch, 115, 118
– econometrisch, 104

- bij simulatie, 248
- statisch, 115
- volledig, 106
- vs. werkelijkheid, 116
MODIGLIANI, F., 278, 306
MONHEMIUS, W., 305
monteur(s) (reparateurs), 242, 243
- optimale aantal, 243-245, 248-257
MOOD, A. M., 340
MORGENSTERN, O., 203, 226
MORONEY, M. J., 180, 340
morra, 206-208
MORSE, P. M., 245
MOUCHART, M., 154, 155
MUTH, J. F., 278, 306

Nachtegaal, 224, 225
NANUS, B., 267
Nederlandsche Economische Hoogeschool, 11
netwerk, 70, 71
NEUMANN, J. VON, 203, 226
nevenvoorwaarde(n), 16, 18-25, 27-29, 50-69, 75
- prijs van, 33, 35
niet-negatieve oplossing, 32
niet-negativiteitsrestricties, 16, 17, 20, 21, 23
niet-nulspelen, 224, 225
noodzakelijk goed, 351
noordwesthoekregel, 53-55, 63
Northwest corner rule, 53
noten, ink. el., 358
nulspelen, 205
nutsfunctie, 353

Oligopolistische markt, 260
omschakelingskosten, 269, 282-285 (*zie ook* instelkosten)
onafhankelijkheid van kansvariabelen, 184
onbekenden, 15
onderlinge leveringen, 83
- voorspelling van, 148, 154
onderschatting, 133-139
- van veranderingen, 145-147
ongelijkheid, 16-19, 28
ontoelaatbare waarden, 22
onzekerheid, 157
oorlogsspelen, 258
opbrengst, verwachte, *zie onder* verwachting
operationele research, 10
Operations Research, 10
oplossing,
- van lineaire vergelijkingen, 31, 92, 93
- niet-negatieve, 32, 53
oplossing (optimale),
- van lineair programmeringsprobleem, 25, 27, 30-33, 57-67, 69
oplossingsmethode, 48
- grafische, 19-30
- algebraïsch-numerieke, 30-33
oplossingstechniek, 19-33, 47, 48, 75
overheid, 77, 81, 102
overheidsuitgaven, effect op betalingsbalans, 122, 123
overschatting, 133
- van veranderingen, 145-147
overwerk, kosten van, 281, 282

Pad, kritieke, 69-76
- langste, 71
PALM, C., 245
PANNE, C. VAN DE, 143, 154, 155, 306
papier(nijverheid), 82-84, 95-97
patates frites, ink. el., 358
pay-off matrix, 193
personeel, aantrekken en ontslaan, 279-281
PERT, 69, 76
Pert-o-graph, 71
Pittsburgh Plate Glass Company, 278

Plan, Centraal Economisch, 120, 121
– voorspellingskwaliteit van, 142-147
Planbureau, Centraal, 105, 120-125, 140
Polarisraket, 70
politiek, macro-economische, 112, 113, 137
populatie, 320
product, nationaal, 105
productie, 18-35, 40-47
– totale, 80-88
productiecentra, 52
productiekosten, 41
productiemiddelen, 99
productieprogramma, 20, 25
productiestatistieken, 79
Program Evaluation and Review Technique, 69
programmeren,
– met gehele getallen, 48, 52
– kwadratisch, 48
– lineair, 14-17, 19, 47, 48, 75, 76, 223n.
prijs,
– van nevenvoorwaarde, 33-35
– van restrictie, 35
– van voedingsmiddelen, 18
prijselasticiteit, 350, 353-355
prijsindex, 109, 346, 347
prijsniveau, 346, 347
puntvoorspellingen, 128
pijlenschema, 113, 114

Quiz, 67

Radio's, 18-27
radiofabrikant, 18
RAIFFA, H., 203, 226
RAND Corporation, 251
rechte lijn(en), 22, 23
– snijpunt, 31
– evenwijdige, 31
– samenvallende, 31
reëel inkomen, 345, 346
regressiecoëfficiënt, 318
regressielijn, 318
regressievergelijking, 318
REITH, P. F., 76
rekenmachines, electronische, 9, 71, 94, 251, 261, 362
relatieve prijzen, 345, 348
reparateurs, *zie* monteurs
restrictie, 16
– arbeids-, 18, 19, 29
– capaciteits-, 19, 23
– grafische voorstelling, 21, 22
– kosten van, 34, 35
– niet-negativiteits-, 16-19, 20, 21, 23
– prijs van, 35
REY, G., 147, 154, 155
RHENMAN, E., 267
richting, foutief voorspelde, 145
risico, 39
rondedans van plussen en minnen, 65, 66
room, ink. el., 358
rundvlees, ink. el., 358
rijlengte (gemiddelde), 235-241

SAATY, T. L., 245
samenhang in de economie, 77, 101
schapenvlees, ink. el., 358
schatten, statistisch, 319
– in regressieanalyse, 326-340
schattingen, zuivere, 324-326
schema (van dit boek), 12
Schiphol, 230
SCHLAIFER, R., 203
schoeisel, ink. el., 352
schoeisel(industrie), 82-84, 95-97
SCHOUTEN, D. B. J., 140, 154, 155
sectoren, 78-100
sectorinkomen, 80, 81
sectorproductie, 79, 80, 86-93
seizoenpatroon, 40

seriegrootte, optimale, 268-277, 305
sigarenindustrie, werkloosheid in, 126, 344
significant, 355
SIMON, H. A., 278, 305, 306
simplex-techniek, 30-33, 47, 48, 49, 52
simulatie,
- algemeen, 248
- bij voetbalpool, 246-248
- bij wachttijdproblemen, 248-254
SISSON, R. L., 267
SITTIG, J., 245
SMITH, ADAM, 77
Society, Econometric, 10
spel(en),
- algemeen, 208
- van één persoon, 188
- van twee personen, 204-224
- van meer personen, 224, 225
- boter, kaas, en eieren, 208, 209
- kruis of munt, 206, 214-219
- morra, 207, 208
- waarde van, 212
spelingsvariabelen, 27-30, 35, 51, 52
spoorboekje, 123
spreiding, 168
standaarddeviatie, 169
- van bedieningstijden, 233
- van tussenaankomsttijden, 232
standaardfout (van regressieco-ëfficiënt), 334, 340-343
Stastok, 228, 229
stationnaire toestand, 236
Statistiek, Centraal Bureau voor de, *zie* C.B.S.
Statistiek, Vereniging voor, 154
steekproef, 320
steekproefgemiddelde, verwachting van, 325
stelsel van vergelijkingen 32
Stichting Studiecentrum voor Administratieve Automatisering, 9
STIGLER, G. J., 14, 47, 48
STILIAN, G. N., 76
STONE, R., 356, 357, 363
STRAELEN, R. A. VAN, 100
strategie, 185
- in boter, kaas en eieren, 209
- dominerende, 197, 205
- gemengde, 214-217
- in schaakspel, 185-187
- voorwaarden voor, 187
structuurvergelijkingen, 107

Tabaksproducten, 82-84, 95-97
tabakswaren, ink. el., 352
tankschepen, 52
taptemelk, ink. el., 358
tendentie, centrale, 168
textiel, ink. el., 352
textiel(nijverheid), 82-85, 95-98
thee, ink. el., 358, 361
THEIL, H., 124, 125, 154, 155, 203, 305, 306, 363, 364
TILANUS, C. B., 147, 154, 155
TINBERGEN, J., 105, 124
toelaatbaar gebied, 21, 24, 25, 27
transactiebedragen, 79
transactieschema, 78-81
transportplan, 53-67
transportprobleem, 48-68, 75, 76
trekkingsprocedure, aselecte, 323
tussenaankomsttijden, 230

Uitgaven, consumptieve, 94
zie ook consument

VAJDA, S., 48, 75
variabele,
- algemeen, 10
- beheerste, 16, 17
- beslissings-, 17, 27-30
- endogene, 106-118
- exogene, 106-119

– lopende, 112-119
– nul-, 29, 32, 66
– positieve, 29, 30, 32, 48, 53, 66, 69
– spelings-, 27-30
– vertraagde, 112-118
variantie, 168, 169
variatiecoëfficiënt, 233
varkensfokker, 224, 225
varkensvlees, ink. el., 358
VENEKAMP, P. E., 100
verdeling,
– van bedieningstijd, 233, 234
– continue, 172, 171-180, 229
– discrete, 164-171, 229, 336
– exponentiële, 232, 242
– marginale, 182
– normale, 175-180
– van toekomstige afzet, 290
– van tussenaankomsttijd, 230-232
verdelingsfunctie, 165, 166, 179
verffabriek, 278
vergelijking, lineaire, 10, 22-24
– van de eerste graad, 10
– van de tweede graad, 10
vergelijkingenstelsel, 32
– econometrisch, 130
verifieerbaarheid, 127
verlies- en winstrekening, 111
vertraging, 108, 109, 112, 113
verwachting, mathematische, 167-171
– van bedieningstijd, 233
– van exponentiële verdeling, 232
– van marginale verdeling, 183
– van opbrengst, 194-202, 216-224
– van steekproefgemiddelde, 325
– van totale kosten, 288
– van tussenaankomsttijd, 232
– van winst, 194-202, 216-224
verwarming, ink. el., 352
VIRENQUE, P. H., 100
vis, ink. el., 358
vis, gebakken, ink. el., 358
vis- en visconserven, ink. el., 352
vlak, twee helften van, 23
vlees, ink. el., 361
vlees in blik, ink. el., 358
vleeswaren. ink. el., 361
vliegtuigfabriek, 188-203
voedingseisen, 14
voedingspakket, kosten van, 15-17
volkshuishouding, 81
voorkeursfuncties, kwadratische maatschappelijke, 306
voorraad, 41-47, 268-277
– optimale, 283-285
voorraadkosten, 41-47, 270, 282-285
voorrang, 233, 238-241, 239
voorspellingen,
– economische, 126-131
– interval-, 128
– lijn der volmaakte, 133
– onvoorwaardelijke, 126, 127, 142-147
– punt-, 128
– van toekomstige beslissingen, 301-305
– van voetbaluitslagen, 146
– voorwaardelijke, 126, 127, 140-142, 147
– zuivere, 302
voorspellingsprocedures, 129, 130
– aan- en afvoertabel, 147-154
– econometrisch model, 130, 139-142
– enquête, 131
– extrapolatie, 130
– naïeve methode, 153, 154
vorm,
– herleide, 107, 116-118
– structurele, 107
VORST, J. I., 154, 155

vraag, 44
– consumptieve, 89
zie ook consument
vraagfunctie, 353-355
vrije wil, 116, 131

Waarde,
– van eenheid capaciteit, 34
– van spel, 212, 223
waarschijnlijkheid, 159-166
waarschijnlijkheidsrekening, 158-184
– axioma's, 162
– beginselen, 160-180
– toepassingen, 160, 193-203, 214-224, 229-245 246-257, 288-305, 319-343
waarschijnlijkheidsverdeling, *zie* verdeling
waarschijnlijkheidsuitspraken, 129
wachttijden, 72
wachttijdproblemen,
– inleiding, 227-229
– simulatie, 248-254
– theorie, 228-237
water, ink. el., 352
wegen, 167
werkgelegenheid, 101
werkloosheid, 101-104, 126
WHITIN, T. M., 305, 306
WILLIAMS, J. D., 226
winst, 18, 36, 80
– maximale, 25, 27
– totale, 18, 24
– verwachte, *zie* verwachting
winstinkomen, 108-111
winstlijn, 24, 26
wiskunde, toepassingsgebieden, 9
WOLFF, P. DE, 340
woninginrichting, ink. el., 352
WOOD, M. K., 47
WIJVEKATE, M. L., 180, 340

IJKEL, J., 267

Zadelpunt, 212, 213, 222
zekerheidsequivalentie, 291
zet, 208
zuivelprodukten, ink. el., 352

AULA

het wetenschappelijke pocketboek

A 3 **Over de pijn** *Prof. dr. F. J. J. Buytendijk*
Een veelzijdige benadering van het alomvattende probleem van de pijn.

A 4 **Cybernetica** *Prof. dr. S. T. Bok*
Hoe sturen wij ons leven, ons werk en onze machines? Een inleiding tot een zeer actuele tak van wetenschap.

A 6 **Etymologisch woordenboek** *Dr. J. de Vries*
Waar komen onze woorden vandaan? De betekenis van het woord en zijn herkomst.

A 13 **De arbeidende klasse in Nederland in de 19e eeuw**
Prof. dr. I. J. Brugmans
Een diepgaande studie van de economische en sociale verhoudingen tussen 1813 en 1870.

A 14 **Moderne economie** *Prof. dr. J. Pen*
Wat bepaalt het nationaal inkomen, de welvaart en de werkgelegenheid?
Duidelijke uiteenzetting voor de leek.

A 21 **De zeventiende eeuw** *Sir George Clark*
Een veelzijdig beeld van een boeiend tijdvak: bevolking, politiek, economie, wetenschap, kunst enz.

A 22 **Geschiedenis der muziek** *Curt Sachs*
Een van de bekendste musicologen geeft een meesterlijke beschrijving van de wording en ontwikkeling der muziek.

A 24 **Latijns-Nederlands woordenboek** *Drs. H. H. Mallinckrodt*

A 25 **Heldenlied en heldensage** *Dr. J. de Vries*
Algemene inleiding op het heldenepos, aan de hand van de epiek der klassieken.

A 32 **De agrarische geschiedenis van West-Europa (500-1850)**
Prof. dr. B. H. Slicher van Bath
Een sociaal-economische geschiedenis van West-Europa.

A 35 **Het verschijnsel mens** *Pierre Teilhard de Chardin*
Een nieuw licht op de evolutieleer door de vermaarde paleontoloog.

A 37 **Natuurwetenschap en techniek** *Prof. dr. A. G. M. van Melsen*
Een wijsgerige bezinning. De aard van de natuurwetenschap en de invloed op de beschaving.

A 39 **Verklarende statistiek** *M. L. Wijvekate*
Het onderscheiden van toeval, schijn en werkelijkheid in cijfermateriaal.

A 40 **Noord en Zuid** *Prof. dr. P. Geyl*
Eenheid en tweeheid in de Lage Landen. Studies over de Groot-Nederlandse gedachte.

A 45 **Karl Marx** *Prof. dr. W. Banning*
Leven, leer en betekenis. De ontwikkeling van een levensbeschouwing; de vorm van het marxisme in de huidige tijd.

A 47 **Spaans-Nederlands woordenboek** *Dr. S. A. Vosters*

A 48 **Nederlands-Spaans woordenboek** *Dr. S. A. Vosters*

A 51 **Evolutie en erfelijkheid** *Th. Dobzhansky*
Een benadering van het evolutieprobleem op basis der genetica; van Darwin tot heden. Geïllustreerd.

AULA
het wetenschappelijke pocketboek

A 52 **Sociale psychologie in drie dimensies** *J. A. A. van Leent*
De verhouding tot de sociologie in de richting van drie hoofdstromingen.

A 53 **De ontcijfering van Griekenlands oudste schrift** *John Chadwick*
De ontraadseling van het oud-Kretensisch schrift, het zg. Lineair B.

A 54 **Schriftelijk rapporteren** *Dr. ir. H. de Boer e.a.*
Een praktische handleiding bij het samenstellen van rapporten, nota's, memoranda, scripties en dergelijke.

A 57 **Inleiding tot de wetenschap der prehistorie** *Hans Jürgen Eggers*
Methoden en problemen van prehistorie en archeologie; de vaststelling van de relatieve en absolute ouderdom. Geïll.

A 58 **Psychologie der beroepskeuze** *Dr. S. Wiegersma*
Classificatie en analyse van de diverse beroepen; de psychologische, lichamelijke en sociaal-economische factoren.

A 59 **Zeventiende-eeuwse geestesbloei** *Dr. Cornelia W. Roldanus*
Een karakteristiek van de voornaamste geestelijke kenmerken van Nederlands Gouden Eeuw. Geïllustreerd.

A 60 **Geschiedenis der gewervelde dieren** *Dr. Emil Kuhn-Schnyder*
Het evolutieverloop met zijn beslissende fasen en dramatische keerpunten. Geïllustreerd.

A 61 **De vroegste geschiedenis van de mens** *Rudolf Grahmann*
De halfmens, de oermens, en de prehistorische mens: zijn biologische en culturele ontwikkeling. Geïllustreerd.

A 62 **Middeleeuwse kunst in overgang** *H. Jantzen*
Het tijdperk der Saksische keizers. Kunsthistorische aspecten van het Duitse keizerrijk van 936-1002. Geïllustreerd.

A 63 **Christelijke theologie van het Oude Testament**
George A. F. Knight
Een ontdekkingstocht door de bijbel; een afgerond overzicht van de gehele oudtestamentische theologie.

A 64 **Algemene zoölogie** *Alfred Kühn*
Een systematische behandeling van de vormleer, de fysiologie en de dierlijke organismen. Geïllustreerd.

A 65 **Amerika in beweging** *R. L. Bruckberger*
De industriële en sociale revolutie van de 20ste eeuw. Uitvoerige analyse van de Amerikaanse beschaving.

A 66 **Inleiding tot het Nieuwe Testament** *Dr. A. F. J. Klijn*
Het Nieuwe Testament, gezien in het licht van de apostolische verkondiging.

A 68 **Existentiële fenomenologie** *Dr. W. Luijpen*
Een zelfstandige wijsgerige verhandeling die nauw aansluit bij de grote figuren van de hedendaagse filosofie.

A 69 **Het ontstaan der beschaving in het Nabije Oosten**
Prof. dr. Henri Frankfort
Een knappe samenvatting en boeiende inleiding over de oudste culturen in Mesopotamië en Egypte.

A 71 **De geestelijke evolutie van de mens** *Gardner Murphy*
De problemen en mogelijkheden van de Westerse wereld; de geciviliseerde mens op weg naar de toekomst.

AULA
het wetenschappelijke pocketboek

A 72 **Keerpunten in de fysica** *R. J. Blin-Stoyle* e.a.
In een aantal voordrachten geven hoogleraren uit Oxford hun visie over de veranderingen in de fysische ideeën.

A 73 **Logisch denken** *E. R. Emmet*
Begrippen en methoden van een elementaire logica; met een kort overzicht van de moderne, symbolische logistiek.

A 74 **De bedreigde mens** *Hans Zbinden*
De geestelijke en sociale situatie van deze tijd. Over de moderne levensangst, de democratie, de positie van het kind, en hiermee samenhangende problemen.

A 75 **Het Noorden op weg naar zelfstandigheid** *Robert Fruin*
Drie opstellen van de eminente Leidse historicus over de Tachtigjarige Oorlog.

A 77 **De ontvoering van Europa** *Luis Diez del Corral*
Een historische interpretatie van onze tijd. De culturele, sociale en politieke situatie van een bedreigd Europa.

A 78 **Wijsgerige methoden in de moderne wetenschap**
J. M. Bochénski
Een voortreffelijke inleiding tot de elementen en systemen van het logische denken.

A 79 **Moderne filosofie** *Ludwig Landgrebe*
Een duidelijke wegwijzer in de gecompliceerde wereld van de wijsbegeerte: de ethische, ontologische en theologische aspecten.

A 80 **Het aardige van de economie** *Prof. dr. J. Pen*
Studies en beschouwingen over verschillende facetten van het economische en sociale leven.

A 81 **De gestalte der toekomst** *Romano Guardini*
Het bekende essay, aangevuld met twee verhandelingen over: de moderne mens en het probleem van de macht.

A 82 **Keerpunt der Middeleeuwen** *F. v. d. Meer*
Een groots overzicht van de gotische bouwkunst in Frankrijk. Geïllustreerd.

A 83 **Het sociale leven der dieren** *A. Portmann*
De psychologie der dieren; de organisatievormen van hogere diersoorten. Geïllustreerd.

A 84 **Mythologie der Grieken** *Georges Méautis*
Het Olympische godendom, de goden van de oertijden en de heros-figuren.

A 85 **Woordenboek der Noord- en Zuidnederlandse plaatsnamen**
Dr. J. de Vries
De oorsprong en betekenis van de namen van dorpen en steden.

A 87 **Het Thomas-evangelie** *Robert M. Grant* en *David N. Freedman*
Vertaling en toelichting op deze verzameling 'woorden en uitspraken' van Jezus.

A 88 **De problemen der hedendaagse erotiek** *Charlotte Köhn-Behrens*
De karaktereigenschappen van man en vrouw: de verschillende aspecten van de verhouding der geslachten.

AULA

het wetenschappelijke pocketboek

A 89 **Sociale antropologie** *Raymond Firth*
De structuren van de verschillende samenlevingen; de wijze waarop vreemde en soms primitieve volken hun leven organiseren.

A 90 **Luther in leven en werk** *Gerhard Ritter*
Een beeld van de grote kerkhervormer tegen de achtergrond van het Duitse leven in de 16e eeuw.

A 93 **Heldensagen** *Dr. J. de Vries*
Een kennismaking met een twintigtal sagen en liederen, waaronder: Siegfried en de Nibelungen, het Kudrun-epos, Beowulf en het lied van de Hunnenslag.

A 94 **De betekenis van de evolutie** *George Gaylord Simpson*
Het evolutieproces, de plaats van de mens in deze miljardjarige ontwikkeling.

A 95 **De fenomenologie van Merleau-Ponty** *Prof. dr. R. C. Kwant*
Zijn standpunt tegenover en naast belangrijke figuren uit de hedendaagse Franse wijsbegeerte.

A 96 **Algemene volkenkunde** *Kunz Dittmer*
Een inleiding tot de geschiedenis en wetenschap der volkenkunde. Vorm, gestalte en ontwikkeling tot hogere culturen.

A 97 **De geschiedenis van het begrip atoom**
Prof. dr. A. G. M. van Melsen
Van atomos naar atoom. De moderne atoomtheorie en de samenhang met de nieuwste stromingen.

A 98 **Beschouwingen over Europese kunst** *Eric Newton*
De Italiaanse en Duitse renaissance, de 17e-eeuwse Nederlandse school. Geïllustreerd.

A 99 **Algemene celleer** *Joseph G. Hoffman*
De bouwstenen van het leven. De cel als basis-eenheid van het leven: bouw, ontwikkeling en organisatie.

A 100 **Geschiedenis van de kerk in Nederland**
Dr. A. G. Weiler, Prof. dr. L. J. Rogier, dr. O. J. de Jong, Prof. C. W. Mönnich
Een uniek boek over de kerkelijke stromingen door de eeuwen heen.

A 101 **Radio-astronomie** *F. Graham Smith*
Een overzicht van de vorderingen op het terrein van de radio-astronomie vanaf het ontstaan in 1932.

A 102 **Geschiedenis van Egypte** *Eberhard Otto*
Het rijk der farao's vanaf 3000 v.Chr. tot aan de ondergang in de dagen van Alexander de Grote, 330 v.Chr.

A 103 **Fundamentele problemen der hedendaagse wijsbegeerte**
Gerhard Krüger
Van de mythische, voor-filosofische voorstelling der Grieken tot een beschouwing van de eigentijdse filosofie.

A 105 **De spanning tussen orde en vrijheid** *Alfred von Martin*
Sociologische kernproblemen. Een historische en systematische verantwoording.

AULA

het wetenschappelijk pocketboek

A 106 **Het unieke van de mens** *R. J. Harrison*
Een zuiver biologische benadering van het ingewikkelde complex dat mens heet: lichaamsstructuur, orgaanfuncties, typische kenmerken. Met vele illustraties.

A 107 **Het trieste der tropen** *Claude Levi-Strauss*
De problematiek van de beschaving belicht aan de hand van een antropologische studie.

A 108 **De gezondheid van het kind** *Richard W. B. Ellis*
Het kind vanaf de conceptie tot de puberteit, de invloeden van erfelijkheid en milieu. Geïllustreerd.

A 109 **Algemene prehistorie** *Grahame Clark*
Een overzicht van de voorgeschiedenis der gehele wereld en van de ontwikkeling van de mens: zijn verspreiding over de aarde, de groei van de beschaving. Geïllustreerd.

A 110 **Psychologie van de droom** *J. A. Hadfield*
Een beschouwing over droom en uitleg; de theoretische ontwikkeling sinds Freud en het belang voor de moderne mens.

A 111 **Liefde in het oude Griekenland** *Robert Flacelière*
De verhouding tussen man en vrouw; de achtergronden en ontwikkeling van het liefdeleven vanaf de tijd van Homeros.

A 112 **Inleiding tot de meteorologie** *Dr. F. H. Schmidt*
De wetenschap van het weer en de methoden der meteorologen: grondbegrippen en praktijk van de weersvoorspelling.

A 113 **Hedendaagse psychiatrie** *David Stafford-Clark*
De ontwikkeling van de psychiatrie, de grenzen van het normale en abnormale.

A 114 **Causaliteit en waarschijnlijkheid in de moderne fysica**
David Bohm
Kritische bespreking van de quantentheorie: herwaardering van de causaliteit naar ideeën van Einstein.

A 115 **Het einde van de antieke wereld** *Santo Mazzarino*
Het verval van staten en beschavingen, weerspiegeld in de opvatting van historici, van Polybius tot Toynbee.

A 116 **De zin der geschiedenis** *Raymond Aron*
Acht belangrijke essays over de historie in verband met het leven en de aspecten van de moderne politiek na 1914.

A 117 **De vijf fasen van economische groei** *W. W. Rostow*
Een inleiding tot de studie van de moderne economische geschiedenis.

A 118 **Moderne biologie** *William S. Beck*
Ontwikkeling en vooruitzicht. Een inzicht in de denkwijze en methoden van de hedendaagse biologie.

A 119 **Godsdienst in de prehistorie** *E. O. James*
Het religieuze leven van de prehistorische mens aan de hand van het beschikbare archeologische materiaal benaderd.

A 120 **Inleiding tot de algemene economie** *A. J. Brown*
Logische en overzichtelijke behandeling van actuele, economische problemen van deze tijd.

A 121 **De vroegste geschiedenis van de samenleving** *Grahame Clark*
Doel en wezen van de archeologie. Het culturele, religieuze en economische leven van prehistorische gemeenschappen.

MARKA HET POCKETBOEK VOOR ORGANISATIE EN BEDRIJF

De drie auteurs, geboren te Amsterdam (1924), Semarang (1936) en Leerbroek (1934), werken op het Econometrisch Instituut van de Nederlandsche Economische Hoogeschool te Rotterdam en doceren aldaar in de kwantitatief-economische studierichting. Henri Theil's research lag oorspronkelijk op het terrein van de econometrie in engere zin, vooral toen hij nog aan het Centraal Planbureau verbonden was, maar in de laatste jaren ging de operationele research mede in klimmende mate zijn belangstelling trekken. Er zijn drie boeken van zijn hand: *Linear Aggregation of Economic Relations* (1954), *Economic Forecasts and Policy* (1958, 2e druk 1961), *Optimal Decision Rules for Government and Industry* (1963). Sinds zijn benoeming tot buitengewoon hoogleraar in de econometrie te Rotterdam in 1953 was hij o.m. gasthoogleraar aan de Harvard University. In 1956 werd hij directeur van het toen opgerichte Econometrisch Instituut, waarmee in 1960 het International Center for Management Science geaffilieerd werd. John C. G. Boot's belangstelling gaat in het bijzonder uit naar de theorie en de toepassing van het mathematisch programmeren en naar de speltheorie. Het consumentengedrag en de indexcijfers behoren tot de favoriete onderwerpen van Teun Kloek.

3500060